

# Analyse vectorielle

Antoine Gournay

Institut de Mathématiques,  
Université de Neuchâtel  
Suisse

Décembre, 2011

Notes de Cours





# Table des matières

<b>Introduction</b>	<b>v</b>
<b>1 L'espace euclidien.</b>	<b>1</b>
1.1 Espace vectoriel . . . . .	1
1.2 Produit scalaire . . . . .	3
1.3 Norme . . . . .	4
1.4 Produit vectoriel ou extérieur . . . . .	5
1.4.i Dans $\mathbb{R}^2$ . . . . .	6
1.A Produit extérieur dans $\mathbb{R}^n$ . . . . .	7
1.A.i Puissances extérieures supérieures . . . . .	7
1.B Espaces vectoriels abstraits et axiomes du choix . . . . .	8
<b>2 Applications, limites et continuité</b>	<b>11</b>
2.1 Représentations graphiques des fonctions . . . . .	12
2.2 Topologie dans $\mathbb{R}^n$ . . . . .	14
2.2.i Espace vectoriel vs espace topologique . . . . .	16
2.3 Limites . . . . .	17
2.4 Continuité . . . . .	19
<b>3 Dérivées et dérivabilité</b>	<b>23</b>
3.1 Dérivées partielles . . . . .	24
3.2 Application linéaire et dérivabilité . . . . .	26
3.3 Gradient et matrice de Jacobi . . . . .	29
3.4 Dérivation composée . . . . .	33
3.4.i Un théorème d'Euler sur les fonctions homogènes . . . . .	35
3.5 Dérivées d'ordres supérieurs . . . . .	36
3.6 Interprétation géométrique quand $m = 1$ ou $n = 1$ . . . . .	39
3.6.i Dérivée directionnelle . . . . .	39
3.6.ii Potentiel . . . . .	41
3.6.iii Vitesse . . . . .	41
3.7 Espace(s) tangent(s) . . . . .	42
3.7.i Tangence aux ensembles de niveau . . . . .	42

## TABLE DES MATIÈRES

---

3.7.ii	Espace des vecteurs vitesse . . . . .	46
3.7.iii	Image d'une application linéarisée . . . . .	48
3.7.iv	Premier lien entre les trois définitions . . . . .	50
3.A	Dérivée de Gâteaux . . . . .	54
<b>4</b>	<b>Courbes</b>	<b>57</b>
4.1	Paramétrage . . . . .	58
4.2	Longueur ; cas Lipschitz . . . . .	62
4.3	Longueur ; cas $C^1$ par morceaux . . . . .	65
4.4	Paramétrage par la longueur d'arc . . . . .	71
4.5	Intégrale curviligne et travail . . . . .	76
4.6	Courbure . . . . .	78
4.A	Torsion . . . . .	81
4.B	Théorème fondamental de l'algèbre . . . . .	82
<b>5</b>	<b>Approximation et étude I</b>	<b>85</b>
5.1	Formule de Taylor . . . . .	85
5.2	Points critiques et singuliers . . . . .	89
5.3	Matrice de Hesse ou hessienne . . . . .	92
5.3.i	Lorsque même la dérivée seconde ne suffit plus... . . . .	97
<b>6</b>	<b>Résolution et étude II</b>	<b>99</b>
6.1	Méthode de Newton . . . . .	99
6.2	Théorème des applications implicites . . . . .	103
6.2.i	Calcul de la dérivée d'une fonction implicite . . . . .	106
6.2.ii	Approximation de Taylor pour les fonctions implicites . . . . .	109
6.3	Espace tangent et sous-variétés . . . . .	109
6.3.i	Cas particuliers . . . . .	116
6.4	Multiplicateurs de Lagrange . . . . .	116
6.4.i	Étude de fonction avec le bord du domaine. . . . .	123
6.4.ii	Inégalité arithmétique-géométrique et théorème spectral . . . . .	127
6.4.iii	Forme quadratique du lagrangien . . . . .	128
6.4.iv	Les contraintes via les paramétrages . . . . .	133
6.A	Dimension . . . . .	133
6.B	La conjecture du Jacobien . . . . .	134
<b>7</b>	<b>Intégrales multiples</b>	<b>137</b>
7.1	Pavés . . . . .	137
7.2	Domaines plus généraux . . . . .	142
7.3	Évaluation . . . . .	143
7.3.i	Intégrales d'applications . . . . .	146
7.4	Changement de variable . . . . .	147

7.A	Intégrales impropres . . . . .	151
7.B	Théorème de la valeur moyenne . . . . .	152
<b>8</b>	<b>Surfaces</b>	<b>155</b>
8.1	Vecteur normal et orientabilité dans $\mathbb{R}^3$ . . . . .	156
8.2	L'orientabilité dans $\mathbb{R}^m$ . . . . .	160
8.3	Surfaces à bord et orientation induite . . . . .	166
8.4	Première forme quadratique : longueur et aire . . . . .	170
8.5	Deuxième forme fondamentale : courbure . . . . .	175
8.A	Le calcul fonctionnel pour les matrices symétriques . . . . .	183
8.B	Combinatoire et courbure négative . . . . .	185
<b>9</b>	<b>Champs de vecteurs</b>	<b>189</b>
9.1	Divergence, rotationnel et laplacien . . . . .	189
9.2	Théorème de Green et de la divergence, cas simple . . . . .	193
9.3	Paramétrage et théorème de Stokes . . . . .	196
9.4	Simple connexité et intégrabilité . . . . .	199
<b>10</b>	<b>Le calcul des variations</b>	<b>201</b>
10.1	Géodésiques et longueur minimale . . . . .	202
10.2	Surfaces minimales et aire minimale . . . . .	204
	<b>Bibliographie</b>	<b>205</b>

## TABLE DES MATIÈRES

---

# Introduction

La présente version des notes n'est pas encore une version achevée (portions incomplètes, quelques incohérences dues à des changements de notations/définitions, des erreurs qui survivent toujours). Dessins et exemples viennent parfois à manquer ; l'index est incomplet et il faudrait probablement raccourcir certains passages (ou en mettre en appendice). L'honorable lecteur est ainsi prié de transmettre à l'auteur les erreurs qu'il y trouvera, les références qui lui semblent faire défaut dans l'index, et les passages qui lui semblent trop confus ou trop peu illustrés.

Ces notes sont faites pour un cours de  $14 \times 2$  heures. Le matériel contenu dans ces notes dépasse largement 28 heures.

Ce cours succède normalement au[x] premier[s] cours d'analyse qui porte[nt] sur l'étude des applications de la droite réelle dans elle-même. Le sujet en est maintenant les applications de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  (où  $m, n \in \mathbb{Z}_{\geq 1}$ ). En un certain sens, l'analyse de fonctions à plusieurs variables fait apparaître beaucoup de phénomènes nouveaux. D'un côté, comprendre certains sous-ensembles de  $\mathbb{R}^m$  mène éventuellement à la géométrie différentielle. D'un autre, le remplacement adéquat de certains concepts en dimension 1 à la dimension supérieure prépare le (ou invite au) voyage vers l'analyse fonctionnelle. Les motivations pour étudier ces fonctions sont abondantes ; par exemple, dans le modèle de la physique newtonienne, le potentiel est une fonction de  $\mathbb{R}^3$  dans  $\mathbb{R}$  et la force qu'il engendre une fonction de  $\mathbb{R}^3$  dans  $\mathbb{R}^3$ . Il est aussi possible de percevoir les fonctions de  $\mathbb{C}$  dans  $\mathbb{C}$  comme des fonctions de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$  ; ainsi ce cours entretient aussi un lien avec l'analyse complexe.

La plupart des applications seront laissées en exercices, pour des raisons de manque de temps en cours et parce que l'auteur suppose que la curiosité du lecteur le poussera à s'y intéresser de lui-même.

Les buts de ce cours sont

Chapitre 1 : Rappels sur  $\mathbb{R}^m$ , espace vectoriel et topologique.

Chapitre 2 : Rappels d'analyse en plusieurs dimensions : limite et continuité.

Chapitre 3 : Introduire les dérivées (une approximation linéaire d'une application). Les sous-ensembles de  $\mathbb{R}^m$  peuvent aussi être mieux compris en les "linéarisant", ceci mènera à l'introduction du concept de plan tangent.

Chapitre 4 : Un des objets d'intérêt en plusieurs dimensions sont les courbes. Quelques exemples de questions typiques : comment calculer la longueur d'une courbe ? comment faire un pendule dont la période d'oscillation ne dépend pas de l'amplitude ? comment tracer une voie de chemin de fer afin de minimiser la tension sur les rails ? comment faire un arc de voûte de

sorte à minimiser les forces présentes ? Les outils analytiques vont servir pour étudier ces problèmes.

- Chapitre 5 : Des approximations plus fines pour les fonctions sont aussi possibles (formule de Taylor), ce qui permet le début de l'étude de fonction.
- Chapitre 6 : La linéarisation des fonctions permet en quelque sorte de linéariser les équations, et en quelque sorte de résoudre des systèmes d'équations non-linéaires (méthode de Newton, théorème des fonctions inverses, théorème des fonctions implicites). Il se trouve que ceci permet d'unifier les différentes notions de plan tangent (dans certains cas). Comme les sous-ensembles de  $\mathbb{R}^m$  sont subtils, tout cela est en fait nécessaire pour pouvoir faire des études de fonctions lorsque le domaine de définition est contraint.
- Chapitre 7 : L'intégration requiert aussi quelques raffinements par rapport au cas des fonctions réelles à valeurs réelles. En particulier, la formule de changement de variable devient plus subtile. Entre autres, certaines intégrales en une variable ne peuvent être évaluées qu'avec des méthodes de deux variables (la célèbre intégrale de Poisson,  $\int_{-\infty}^{\infty} e^{-x^2} dx$ ).
- Chapitre 8 : Les questions sur les surfaces sont plus subtiles que sur les courbes. Les théorèmes sur les plans tangents permettront de bien caractériser le genre d'objet qui se porte à l'étude. Seules seront ici traitées les questions d'orientabilité, de calcul d'aire et de flux.
- Chapitre 9 : En électromagnétisme, les champs de vecteurs jouent un rôle primordial. Ce chapitre s'appuie sur tous les précédents pour montrer les théorèmes de base qui servent (théorème de la divergence, de Green et de Stokes). Puisqu'il n'est pas donné à tous d'étudier l'électromagnétisme, le lecteur peut se rassurer en sachant que ceux-ci sont aussi utiles en analyse complexe.

Dans ce cours, les sections qui commencent par une lettre ne seront (au mieux) que partiellement traitées en classe ; leur présence vise plus à présenter de la culture générale ou à approfondir certains thèmes abordés. Le texte des preuves ne contrastant pas beaucoup avec le reste, un  $\blacklozenge$  en marque la fin. Pour les mêmes raisons, un  $\star$  termine le texte des définitions, un  $\clubsuit$  clôt les exemples, un  $\spadesuit$  boucle les remarques et les énoncés (lemme, théorème, etc...) sont légèrement encadrés.

Ceci est aussi le moment opportun pour mettre de l'avant quelques bizarreries dans les notations et les concepts de ce texte.

- Une bizarrerie commune vient du fait que les éléments d'un produit cartésien sont écrits comme  $(a, b, c)$  [une ligne] alors que les vecteurs (des éléments de  $\mathbb{R}^n$ , un produit cartésien) sont notés comme des colonnes. Cette bizarrerie est malheureusement entrée dans les moeurs...
- Au chapitre 6, les sous-variétés de  $\mathbb{R}^n$  possèdent déjà une structure différentielle canonique héritée de  $\mathbb{R}^n$ , ainsi il n'y aura pas de références aux cartes (la définition habituelle d'une variété différentielle "abstraite") ; ce concept sera probablement traité dans un cours de géométrie différentielle. C'est l'application inverse d'une carte, appelée ici paramétrage, qui prendra sa place. Cette perspective est plus naturel et accessible (voir par exemple le livre Milnor, "Topology from the differentiable viewpoint"). D'autre part, un théorème difficile



montre qu'une variété lisse est toujours une sous-variété de  $\mathbb{R}^N$ , cette restriction n'est donc pas si grande qu'on pourrait le croire.

- Le chapitre 7 sur l'intégration sera très court. Trois raisons à cela : (a) manque de temps, (b) il est déjà traité dans un cours précédent (calcul différentiel et intégral à plusieurs variables), et (c) il sera de nouveau abordé dans un cours ultérieur (mesure et intégration).
- Au chapitre 8, les partitions de l'unité ne seront pas utilisées (encore une fois, on relègue cette technologie à un cours de géométrie différentielle).

Voici quelques notations qui seront employées tout au long de ce texte :

$\mathbb{Z}$  pour l'ensemble des entiers (positifs, négatifs ou nul).  $\mathbb{R}$  pour l'ensemble réel. Quelques notations qui seront utilisées pour des sous-ensembles de ceux-ci sont  $\mathbb{Z}_{\geq a}$  (des entiers supérieurs ou égaux à  $a$ ),  $\mathbb{R}_{< a}$  ou  $]-\infty, a[$  (des réels strictement inférieurs à  $a$ ),  $]a, b]$  (pour les réels strictement plus grand que  $a$  et plus petits ou égaux à  $b$ ), etc...

Pour  $n \in \mathbb{Z}_{>0}$ ,  $\overline{n}$  signifie l'ensemble  $\{1, 2, 3, \dots, n\}$  des entiers compris entre 1 et  $n$  (inclusivement).

Lorsqu'un symbole qui n'a pas été préalablement défini apparaît dans une égalité  $:=$  c'est qu'il s'agit là de sa définition.

La composition de deux fonctions  $f : A \rightarrow B$  et  $g : B \rightarrow C$ , notée  $g \circ f$ , est la fonction

$$\begin{aligned} g \circ f : A &\rightarrow B \\ x &\mapsto g(f(x)) \end{aligned}$$

Si  $A, B$  sont deux ensembles,  $S \subset B$  un sous-ensemble et  $F : A \rightarrow B$  une application, alors  $F^{-1}(S) := \{a \in A \mid \exists s \in S, F(a) = s\}$  est appelé l'image réciproque de  $S$  par  $F$ .

Si  $U$  est un intervalle ouvert de  $\mathbb{R}$ ,  $k \in \mathbb{Z}_{\geq 0}$  et  $f : U \rightarrow \mathbb{R}$  est une fonction, alors  $f^{[k]}$  est la dérivée d'ordre  $k$  de  $f$ . En particulier,  $f^{[0]} = f$  et  $f^{[1]}$  est plus communément notée  $f'$ .

Finalement, quelques mots sur certaines abréviations :

1. *i.e.* (lat. *id est*) signifie plus ou moins “c'est la même chose que”, “de manière équivalente”, “en d'autres mots”, ...
2. *e.g.* (lat. *exempli gratia*) signifie par exemple “par exemple”.
3. *c'àd.* (fr. c'est-à-dire) est un peu comme *i.e.* mais en moins chic.
4. “ou (équiv.)” pour préciser que la conjonction “ou” qui précède est en fait une équivalence mais que l'auteur n'a pas la motivation nécessaire pour la démontrer.
5. “... X (resp. Y) ... Z (resp. W) ...” signifie qu'on peut lire une première fois la définition/le théorème avec X et Z, puis une seconde fois mais avec Y et W au lieu de X et Z.
6. *mutatis mutandis* = “ce qui devait être changé ayant été changé”, une phrase très utile, comme le lecteur peut s'en douter, car il faut deviner ce qui doit être changé.
7. *ceteris paribus* [*sic stantibus*] = “toute chose égale par ailleurs”, pour bien spécifier qu'on change une quantité, mais que toutes les autres restent égales.

L'auteur aimerait remercier Béatrice de Tilière qui lui a donné les notes de cours des années précédentes, notes dont il s'est inspiré. Il tient à mentionner que les citations en début de chapitre sont sans but subversif.



# Chapitre 1

## L'espace euclidien.

*“Il n’y a pas de définition canonique de canonique”*

-ANONYME

Le but de ce chapitre est d’introduire/de rappeler plusieurs structures sur l’ensemble  $\mathbb{R}^n$  qui, étant les plus communément employées, ne seront probablement pas étrangères au lecteur. Lorsque  $\mathbb{R}^n$  est équipé de toute la panoplie (plus ou moins canonique) qui sera introduite dans ce chapitre, il est alors appelé l’espace euclidien de dimension  $n$  (parfois noté  $\mathbb{E}^n$ ).

L’ensemble  $\mathbb{R}^n$  est l’ensemble des  $n$ -uplets de nombres réels. Les éléments de  $\mathbb{R}^n$  seront écrits en gras, tandis que ceux de  $\mathbb{R}$  en italique (police usuelle dans les formules mathématiques). Ainsi  $\mathbf{x} \in \mathbb{R}^n$  est la donnée de  $n$  nombres réels :  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ , les  $x_i$  ( $i \in \overline{n}$ ) sont appelées les coordonnées du point  $\mathbf{x}$ .

La présence d’un  $^T$  après la parenthèse vient du fait que les vecteurs sont traditionnellement des colonnes. Ainsi, il faut prendre la transposée de la matrice de taille  $1 \times n$  [ou vecteur ligne] pour obtenir une matrice [ou vecteur colonne ou vecteur tout court]. Ce  $^T$  apparaîtra souvent, car il est rarement commode de caser une colonne dans un texte...

Comme dans tous les produits cartésiens, il y a des projections canoniques : soit  $i \in \overline{n}$ , alors la  $i^{\text{ème}}$  projection canonique est définie par

$$\begin{aligned} \pi_i : \quad \mathbb{R}^n &\rightarrow \mathbb{R} \\ (x_1, x_2, \dots, x_n)^T &\mapsto x_i \end{aligned}$$

### 1.1 Espace vectoriel

$\mathbb{R}^n$  peut être muni de plusieurs opérations. Ici, il sera muni seulement de la somme vectorielle et de la multiplication par un scalaire.

- soit  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  et  $\mathbf{y} = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$  alors la somme vectorielle de  $\mathbf{x}$  et  $\mathbf{y}$ , notée  $\mathbf{x} + \mathbf{y}$ , est définie par  $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)^T$ .
- soit  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$  et  $a \in \mathbb{R}$ , la multiplication de  $\mathbf{x}$  par le scalaire  $a$  est  $a\mathbf{x} = (ax_1, ax_2, \dots, ax_n)^T$ .

## 1.1 - Espace vectoriel

---

La somme vectorielle est plus simplement appelée somme (lorsque  $n = 1$  il n'y a en effet pas de différence) et les éléments de  $\mathbb{R}$  sont souvent nommés scalaires (la multiplication scalaire est un changement d'échelle). Deux notations naturelles qui seront employées dans ce texte sont  $\mathbf{0} = (0, 0, \dots, 0)$  (qui sera parfois appelé[e] l'origine ou l'élément neutre) et  $-\mathbf{x} = (-x_1, -x_2, \dots, -x_n)$ .

Les deux opérations ci-dessus font de  $\mathbb{R}^n$  un espace vectoriel, *c'ad.* que les propriétés suivantes sont vérifiées : soient  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$  et  $a, b \in \mathbb{R}$  alors

$$\text{EV1. } (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}) \quad (\text{associativité de la somme})$$

$$\text{EV2. } \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \quad (\text{commutativité})$$

$$\text{EV3. } a(b\mathbf{x}) = (ab)\mathbf{x} \quad (\text{associativité du produit})$$

$$\text{EV4. } a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y} \quad (\text{distributivité vectorielle})$$

$$\text{EV5. } (a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x} \quad (\text{distributivité scalaire})$$

$$\text{EV6. } \mathbf{x} + \mathbf{0} = \mathbf{x} \quad (\text{existence d'un l'élément neutre pour la somme})$$

$$\text{EV7. } \mathbf{x} + (-\mathbf{x}) = \mathbf{0} \quad (\text{existence d'un inverse pour la somme})$$

$$\text{EV8. } 1\mathbf{x} = \mathbf{x} \quad (\text{existence d'un élément neutre pour le produit})$$

La vérification de ces propriétés est une conséquence directe des propriétés de  $\mathbb{R}$  et des définitions utilisées.

Il existe une autre méthode courante pour écrire le vecteur  $\mathbf{x} \in \mathbb{R}^n$ . En effet, soit

$$\mathbf{e}_1 = (1, 0, \dots, 0), \mathbf{e}_2 = (0, 1, \dots, 0), \dots, \mathbf{e}_n = (0, 0, \dots, 1)$$

(autrement dit,  $\forall i, j \in \overline{n}, i \neq j \Rightarrow \pi_j \mathbf{e}_i = 0$  et  $\pi_i \mathbf{e}_i = 1$ ). L'ensemble  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  est appelé base canonique de  $\mathbb{R}^n$ . Tout élément  $\mathbf{x} \in \mathbb{R}^n$  s'exprime grâce à cette base comme

$$(1.1.1) \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n x_i \mathbf{e}_i.$$

Si  $S \subset \mathbb{R}^n$  et  $t \in \mathbb{R}$ , l'expression  $tS$  signifie l'ensemble  $\{t\mathbf{x} \mid \mathbf{x} \in S\}$ .

Finalement, la distinction suivante sera (dans la mesure du possible) faite :

1. le mot élément ou point désignera préférablement un élément de  $\mathbb{R}^n$
2. le terme vecteur aura plutôt un sens vieillot, *c'ad.* qu'un vecteur  $\mathbf{x}$  basé  $\mathbf{a}$ , c'est la donnée d'une direction et d'une longueur (un élément de  $\mathbf{x}$  de  $\mathbb{R}^n$ ) et d'un point de base (un point  $\mathbf{a}$  de  $\mathbb{R}^n$ ), ou, autrement dit, il représente la flèche qui débute en  $\mathbf{a}$  et se termine en  $\mathbf{x} + \mathbf{a}$ .

En particulier, il n'y a aucune différence entre un élément ou point  $\mathbf{x}$  de  $\mathbb{R}^n$  et un vecteur  $\mathbf{x}$  basé en  $\mathbf{0}$  et lorsque le point de base d'un vecteur n'est pas spécifié c'est qu'il s'agit de  $\mathbf{0}$ . Il est évident que, dans la littérature en général, le mot vecteur désigne un élément d'un espace vectoriel. L'emploi de "vecteur" sera fait pour tenter d'insister sur l'aspect géométrique (direction, force, etc...) du concept.

EXERCICE 1: Soit  $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$  et  $\mathbf{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  deux points de  $\mathbb{R}^2$ . Les représenter graphiquement ainsi que les points  $\mathbf{x} + \mathbf{y}$  et  $\mathbf{x} - \mathbf{y}$ .

EXERCICE 2: Soit  $V = \{\mathbf{v} \in \mathbb{R}^3 \mid \mathbf{v} \cdot (2, -1, 3)^T = 0\}$ . Montrer que  $V$  est un sous-espace vectoriel de  $\mathbb{R}^3$  et en trouver une base.

## 1.2 Produit scalaire

Soit  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , le produit scalaire usuel de  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  et  $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ , noté  $\mathbf{x} \cdot \mathbf{y}$  ou encore  $\langle \mathbf{x} | \mathbf{y} \rangle$ , est

$$\mathbf{x} \cdot \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i \in \mathbb{R}.$$

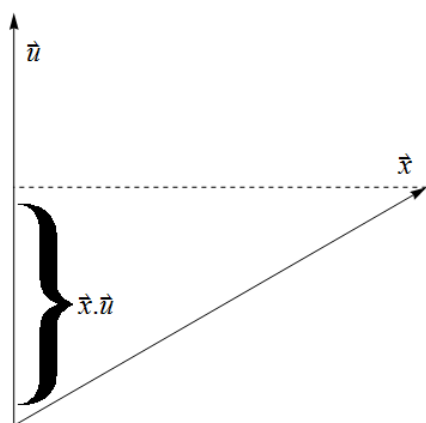
Encore une fois, de cette définition et des propriétés de  $\mathbb{R}$ , il est simple de montrer que le produit scalaire possède les propriétés suivantes : soit  $\mathbf{z} \in \mathbb{R}^n$  et  $a, b \in \mathbb{R}$ , alors

PS1.  $\mathbf{x} \cdot \mathbf{x} \geq 0$  et  $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$  (défini positif)

PS2.  $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$  (commutativité)

PS3.  $(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a\mathbf{x} \cdot \mathbf{z} + b\mathbf{y} \cdot \mathbf{z}$  (linéarité)

[La bilinéarité, c'ad. la propriété de linéarité sur les deux termes, est ici conséquence de la linéarité et de la symétrie.] Deux vecteurs  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  tels que  $\mathbf{x} \cdot \mathbf{y} = 0$  sont dits orthogonaux.



Lorsqu'un des deux vecteurs est unitaire (*i.e.*  $\mathbf{u} \cdot \mathbf{u} = 1$ ), alors il est possible d'interpréter le produit scalaire  $\mathbf{x} \cdot \mathbf{u}$  comme la longueur de  $\mathbf{x}$  dans la direction de  $\mathbf{u}$  (voir dessin ci-contre).

**Remarque 1.2.1.** Si les éléments de  $\mathbb{R}^n$  sont écrits comme des matrice colonnes (*i.e.* de taille  $n \times 1$ ) alors  $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x}$  où dans les termes de droite le produit matriciel est fait et  $\mathbf{y}^T$  est la matrice ligne (de taille  $1 \times n$ ) transposée de la colonne  $\mathbf{y}$ . ♠

Les propriété PS1 à 3 peuvent se démontrer sans trop de souci depuis les propriétés du produit matriciel. Avant de passer à la suite, il est bon de rappeler en quoi ce produit scalaire est un peu arbitraire. En effet, le lecteur se rappellera peut-être qu'étant donné une matrice définie positive (la définition en est rappelée au chapitre 4, définition 5.3.2) notée  $Q$  de taille  $n \times n$ , la quantité  $\mathbf{x} \cdot_Q \mathbf{y}$  définie par  $\mathbf{x}^T Q \mathbf{y}$  est aussi un produit scalaire (*i.e.* satisfait aussi PS1 à 3). De cette façon, il y a moyen de créer beaucoup de produits scalaires différents.

**EXERCICE 3:** Soit  $\Lambda$  une matrice diagonale dont les coefficients sont tous strictement positifs. Soit  $\cdot_\Lambda : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  l'application définie par  $\mathbf{x} \cdot_\Lambda \mathbf{y} = \mathbf{x}^T \Lambda \mathbf{y}$ . En utilisant les propriétés de  $\mathbb{R}$ , montrer que cette application satisfait :

PS1.  $\mathbf{x} \cdot_\Lambda \mathbf{x} \geq 0$  et  $\mathbf{x} \cdot_\Lambda \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$  (défini positif)

PS2.  $\mathbf{x} \cdot_\Lambda \mathbf{y} = \mathbf{y} \cdot_\Lambda \mathbf{x}$  (commutativité)

PS3.  $(a\mathbf{x} + b\mathbf{y}) \cdot_\Lambda \mathbf{z} = a\mathbf{x} \cdot_\Lambda \mathbf{z} + b\mathbf{y} \cdot_\Lambda \mathbf{z}$  (linéarité)

**EXERCICE 4:** Considérer l'espace des fonctions continues sur  $[-1, 1]$ . Soit  $\langle f | g \rangle = \int_{-1}^1 f(x)g(x)dx$ . Montrer que  $\langle \cdot | \cdot \rangle$  est un produit scalaire. Calculer  $\langle f | f \rangle$  lorsque  $f(x) = x^3$ .

## 1.3 Norme

Le produit scalaire induit une norme ; la norme (induite) d'un vecteur  $\mathbf{x}$  est notée  $\|\mathbf{x}\|$  et se définit par

$$(1.3.1) \quad \|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \in \mathbb{R}_{\geq 0}.$$

Dans le cas  $n = 2$ , cette définition coïncide avec celle de longueur par le biais du théorème de Pythagore. Soit  $\mathbf{x} \in \mathbb{R}^n$  et  $a \in \mathbb{R}$ , alors

$$\text{No1. } \|\mathbf{x}\| \geq 0 \text{ et } \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0} \quad (\text{positivité})$$

$$\text{No2. } \|a\mathbf{x}\| = |a|\|\mathbf{x}\| \quad (\text{multiplicativité})$$

$$\text{No3. } \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad (\text{inégalité du triangle})$$

La troisième propriété est un peu plus difficile à démontrer (elle s'appelle aussi parfois propriété d'Archimède). Elle se représente graphiquement par le fait que dans un triangle le chemin direct entre deux sommets est plus court que celui qui passe par les deux autres cotés. La démonstration sera présentée sous peu, en utilisant l'inégalité de Cauchy-Schwarz :

### **Théorème 1.3.2 (Inégalité de Cauchy-Schwarz)**

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \mathbf{x} \cdot \mathbf{y} \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

*Parenthèse* : une des grandes beautés de l'inégalité de Cauchy-Schwarz est qu'elle se généralise à tous les espaces vectoriels munis d'un produit scalaire (et de la norme induite).

DÉMONSTRATION PREUVE DE L'INÉGALITÉ DE CAUCHY-SCHWARZ: Si  $\mathbf{x}$  ou  $\mathbf{y} = \mathbf{0}$ , l'inégalité est trivialement vérifiée. Sinon, il suffit de regarder la fonction (de  $\mathbb{R}$  vers  $\mathbb{R}$ )

$$P(t) = \|\mathbf{x} + t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{y}t + \|\mathbf{y}\|^2 t^2.$$

Cette fonction est en fait un polynôme de degré 2 en  $t$  qui est positive ou nulle pour tout  $t$ . Son discriminant  $4(\mathbf{x} \cdot \mathbf{y})^2 - 4\|\mathbf{x}\|^2 \|\mathbf{y}\|^2$  est donc négatif ou nul, ce qui permet de conclure. ■

DÉMONSTRATION PREUVE DE L'INÉGALITÉ DU TRIANGLE: Il ne s'agit plus que d'un calcul direct :

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}), &= \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2, \\ &&\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2, \\ &&= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned} \quad \blacksquare$$

Remarquez que cette preuve fournit aussi le résultat suivant : deux vecteurs sont orthogonaux si et seulement si  $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ .

Il est aussi maintenant possible de parler d'angle entre deux vecteurs : soit  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , alors l'angle entre  $\mathbf{x}$  et  $\mathbf{y}$  est

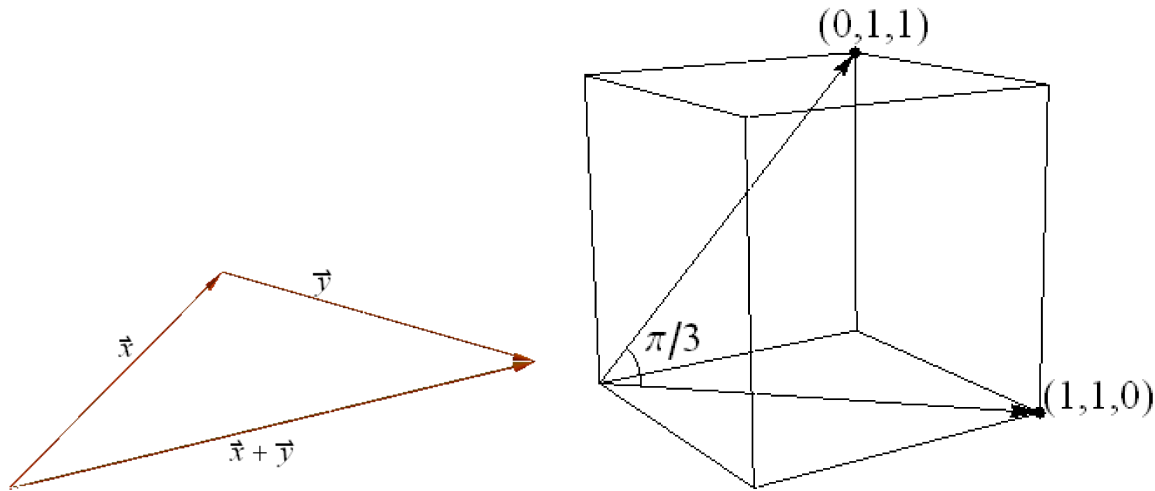
$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \in [0, \pi].$$

La définition est cohérente, car l'inégalité de Cauchy-Schwarz nous assure que l'argument de l'arccos est dans  $[-1, 1]$ . L'angle obtenu est quant à lui compris entre 0 et  $\pi$ . Une fois montrée l'invariance

du produit scalaire lorsque les deux vecteurs subissent une rotation, il est assez facile de voir que cette notion correspond à la notion d'angle usuel.

EXERCICE 5: Soit  $\mathbf{x} = (2, 0, 1)^T$  et  $\mathbf{y} = (1, -1, -1)^T$  deux points de  $\mathbb{R}^3$ . Calculer  $\mathbf{x} + \mathbf{y}$ ,  $\mathbf{x} - \mathbf{y}$ ,  $-2\mathbf{y}$ . Calculer  $\cos \angle(\mathbf{x}, \mathbf{y})$ .

Deux dessins : à gauche celui d'une inégalité du triangle, à droite le calcul d'un angle dans le cube (grâce au produit scalaire)



EXERCICE 6: Calculer

- $\angle((1, 1, 1)^T, (-1, 1, -1)^T)$ ;
- $\cos \angle((1, 2, 3, 4)^T, (4, 3, 2, 1)^T)$ ;
- $\cos \angle((\pi, 0, 1, 0)^T, (1, -1, 1, -1)^T)$ .

EXERCICE 7: Soient  $\mathbf{x}, \mathbf{y}$  deux vecteurs non nuls de  $\mathbb{R}^n$ . Supposons que pour tout  $a \in \mathbb{R}$ ,  $\|\mathbf{x} + a\mathbf{y}\| \geq \|\mathbf{x}\|$ . Montrer que  $\mathbf{x}$  et  $\mathbf{y}$  sont orthogonaux.

## 1.4 Produit vectoriel ou extérieur

Il existe un autre produit dans  $\mathbb{R}^3$  : le produit vectoriel ou extérieur. Il associe à une paire de vecteurs un autre vecteur. Soit  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$  alors le produit extérieur de  $\mathbf{x} = (x_1, x_2, x_3)^T$  et  $\mathbf{y} = (y_1, y_2, y_3)^T$ , noté  $\mathbf{x} \times \mathbf{y}$  ou  $\mathbf{x} \wedge \mathbf{y}$  ou encore  $[\mathbf{x}, \mathbf{y}]$ , est défini par

$$\mathbf{x} \wedge \mathbf{y} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \in \mathbb{R}^3.$$

Voici deux aides-mémoire pour se souvenir de la formule ci-dessus. Le premier consiste à remarquer que dans la  $i^{\text{ème}}$  coordonnée aucun terme en  $x_i$  ou  $y_i$  n'apparaît (et chaque terme contient exactement une coordonnée de  $\mathbf{x}$  et de  $\mathbf{y}$ ). Ensuite, quand un terme est écrit avec le “x” d’abord, son signe est positif si les indices sont dans le “bon ordre” (1, 2, 3, 1) et négatif sinon (1, 3, 2, 1). Le second

provient d'une autre manière plus concise (et plus malpropre) est d'écrire le produit comme

$$\mathbf{x} \wedge \mathbf{y} = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \mathbf{e}_3 = \begin{vmatrix} \mathbf{e}_1 & x_1 & y_1 \\ \mathbf{e}_2 & x_2 & y_2 \\ \mathbf{e}_3 & x_3 & y_3 \end{vmatrix}.$$

Il faut développer le déterminant pour obtenir à nouveau la même expression.

Trois propriétés simples de ce produit sont,  $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$  et  $\forall a, b \in \mathbb{R}$ ,

PV1.  $\mathbf{x} \wedge \mathbf{y} = -\mathbf{y} \wedge \mathbf{x}$  (anti-symétrie)

PV2.  $(a\mathbf{x} + b\mathbf{y}) \wedge \mathbf{z} = a\mathbf{x} \wedge \mathbf{z} + b\mathbf{y} \wedge \mathbf{z}$  (linéarité)

PV3.  $(\mathbf{x} \wedge \mathbf{y}) \cdot \mathbf{x} = 0 = (\mathbf{x} \wedge \mathbf{y}) \cdot \mathbf{y}$  (orthogonalité)

[La bilinéarité est cette fois-ci conséquence de la linéarité et de l'anti-symétrie.] L'anti-symétrie a pour conséquence importante que  $\mathbf{x} \wedge \mathbf{x} = 0$  pour tout  $\mathbf{x}$ . Une autre propriété notable est que, à un signe près,  $\mathbf{x} \wedge \mathbf{y}$  est l'aire du parallélogramme engendré par  $\mathbf{x}$  et  $\mathbf{y}$  (c'est-à-d. celui dont les 4 sommets sont  $0, \mathbf{x}, \mathbf{y}$  et  $\mathbf{x} + \mathbf{y}$ ).

Ainsi, étant donné  $\mathbf{x}, \mathbf{y}$  et  $\mathbf{z} \in \mathbb{R}^3$ ,  $|(\mathbf{x} \wedge \mathbf{y}) \cdot \mathbf{z}|$  est le volume du parallélépipède appuyé sur les trois segments reliant l'origine aux points  $\mathbf{x}, \mathbf{y}$  et  $\mathbf{z}$ . Ceci a pour conséquence que le produit vectoriel de deux vecteurs orthogonaux de norme 1 est un troisième vecteur qui complète la base orthonormale.

L'application  $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \mapsto (\mathbf{x} \wedge \mathbf{y}) \cdot \mathbf{z}$  est parfois nommée produit mixte ou produit scalaire triple. Notons qu'une permutation des trois membres dans le "bon" ordre ne change pas le produit mixte, alors qu'une permutation dans le "mauvais" ordre change son signe, *i.e.*

$$(\mathbf{x} \wedge \mathbf{y}) \cdot \mathbf{z} = (\mathbf{z} \wedge \mathbf{x}) \cdot \mathbf{y} = (\mathbf{y} \wedge \mathbf{z}) \cdot \mathbf{x} = -(\mathbf{x} \wedge \mathbf{z}) \cdot \mathbf{y} = -(\mathbf{y} \wedge \mathbf{x}) \cdot \mathbf{z} = -(\mathbf{z} \wedge \mathbf{y}) \cdot \mathbf{x}.$$

Deux propriétés importantes du produit vectoriel seront encore utilisées.

#### Proposition 1.4.1

Soit  $\mathbf{x}$  et  $\mathbf{y} \in \mathbb{R}^3$  alors

1.  $\|\mathbf{x} \wedge \mathbf{y}\| = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$  où  $\theta \in [0, \pi]$  ;
2.  $\|\mathbf{x} \wedge \mathbf{y}\|^2 + |\mathbf{x} \cdot \mathbf{y}|^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$ .

La démonstration en sera faite (dans le cas général) dans quelques paragraphes.

#### 1.4.i Dans $\mathbb{R}^2$

Dans  $\mathbb{R}^2$ , le produit extérieur permet de définir aussi un produit (qui aura pour résultat un réel). En effet, deux éléments  $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$  de  $\mathbb{R}^2$  peuvent être perçus comme des éléments de  $\mathbb{R}^3$  en ajoutant une coordonnée nulle. Le produit extérieur de  $(x_1, x_2, 0)^T$  et  $(y_1, y_2, 0)^T$  n'aura alors qu'une coordonnée non-nulle, la troisième, dont la valeur sera  $x_1 y_2 - x_2 y_1$ . Cette quantité est (éventuellement à un signe près) l'aire du parallélogramme appuyé sur les deux segments reliant l'origine à  $\mathbf{x}$  et  $\mathbf{y}$ .



## 1.A Produit extérieur dans $\mathbb{R}^n$

Le produit extérieur peut être défini dans  $\mathbb{R}^n$  pour n'importe quel  $n > 1$ . Il n'est cependant pas aussi commode, et surtout le résultat du produit appartient à un espace souvent trop grand,  $\Lambda^2 \mathbb{R}^n$  (heureusement,  $\Lambda^2 \mathbb{R}^2$  s'identifie à  $\mathbb{R}$  et  $\Lambda^2 \mathbb{R}^3$  s'identifie à  $\mathbb{R}^3$ ).

$\Lambda^2 \mathbb{R}^n$  se décrit le plus simplement par sa base (ses éléments seront des combinaisons linéaires des éléments de la base). La base en question sera donnée par les symboles  $\mathbf{e}_i \otimes \mathbf{e}_j$  où  $i$  et  $j \in \bar{n}$  sont distincts, en identifiant cependant  $\mathbf{e}_i \otimes \mathbf{e}_j = -\mathbf{e}_j \otimes \mathbf{e}_i$ . Le produit extérieur est d'abord défini par  $\mathbf{e}_i \wedge \mathbf{e}_j = \mathbf{e}_i \otimes \mathbf{e}_j$ . Par linéarité, il s'étend ensuite à deux vecteurs quelconques de  $\mathbb{R}^n$ . La base de  $\Lambda^2 \mathbb{R}^n$  est de cardinalité  $n(n-1)/2$ , ce qui rend ce produit difficile à réaliser géométriquement si  $n > 3$ .

L'identification  $\Lambda^2 \mathbb{R}^3 \rightarrow \mathbb{R}^3$  est l'application linéaire définie (sur la base) par  $\mathbf{e}_2 \otimes \mathbf{e}_3 \mapsto \mathbf{e}_1$ ,  $\mathbf{e}_1 \otimes \mathbf{e}_3 \mapsto -\mathbf{e}_2$  et  $\mathbf{e}_1 \otimes \mathbf{e}_2 \mapsto \mathbf{e}_3$ .

Il est possible de donner à  $\Lambda^2 \mathbb{R}^n$  un produit scalaire, tout simplement en reprenant le produit canonique mais sur la base  $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i < j \in \bar{n}}$ . Par exemple

$$(3\mathbf{e}_1 \otimes \mathbf{e}_4 - \sqrt{2}\mathbf{e}_4 \otimes \mathbf{e}_2 - 4\mathbf{e}_3 \otimes \mathbf{e}_5) \cdot (\pi\mathbf{e}_1 \otimes \mathbf{e}_4 + 2\mathbf{e}_2 \otimes \mathbf{e}_4 - 2\mathbf{e}_4 \otimes \mathbf{e}_5) = 3\pi + 2\sqrt{2}.$$

À noter qu'il faut s'assurer de l'ordre des composantes (ici  $\mathbf{e}_4 \otimes \mathbf{e}_2 = -\mathbf{e}_2 \otimes \mathbf{e}_4$ ) est toujours le même.

Muni de ce produit scalaire sur  $\Lambda^2 \mathbb{R}^n$  (et de la norme qu'il induit) la proposition 1.4.1 reste vraie.

DÉMONSTRATION DÉMONSTRATION DE LA PROPOSITION 1.4.1: Le premier point est une conséquence du second et de la définition de l'angle.

Pour le second, soit  $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$  et  $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top$ . Il suffit d'écrire

$$\begin{aligned} \|\mathbf{x} \wedge \mathbf{y}\|^2 &= \sum_{0 < i < j \in \bar{n}} (x_i y_j - x_j y_i)^2 = \sum_{0 < i < j \in \bar{n}} (x_i^2 y_j^2 + x_j^2 y_i^2 - 2x_i x_j y_i y_j) = \sum_{i, j \in \bar{n}} x_i^2 y_j^2 - x_i x_j y_i y_j, \\ |\mathbf{x} \cdot \mathbf{y}|^2 &= \left( \sum_{i \in \bar{n}} x_i y_i \right)^2 = \sum_{i, j \in \bar{n}} x_i y_i x_j y_j, \\ \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 &= \left( \sum_{i \in \bar{n}} x_i^2 \right) \left( \sum_{j \in \bar{n}} y_j^2 \right) = \sum_{i, j \in \bar{n}} x_i^2 y_j^2, \end{aligned}$$

et de voir que l'égalité a en effet lieu. Cette démonstration est particulièrement non-géométrique.

Un autre argument est de procéder dans le sens inverse (montrer 1 et en déduire 2). Toutes ces opérations se comportent bien sous l'action des transformations orthogonales (par définition des transformations orthogonales pour  $\cdot$ , il faut travailler un peu plus pour  $\wedge$ ), ce qui fait qu'il est possible de ramener le problème au cas où  $\mathbf{x}$  et  $\mathbf{y}$  sont dans le plan. Alors,  $\|\mathbf{x} \wedge \mathbf{y}\|$  est l'aire du parallélogramme. La hauteur  $h$  du parallélogramme est (en prenant  $\mathbf{y}$  pour la base) est  $h^2 = \|\mathbf{x}\|^2 - |\mathbf{x} \cdot \mathbf{y}|^2 / \|\mathbf{y}\|^2$ . Ainsi, l'aire du parallélogramme est aussi  $\|\mathbf{x}\| \|\mathbf{y}\| \sqrt{1 - \cos^2 \theta}$ . Égaliser ces deux quantités montre la première partie de la proposition. ■

### 1.A.i Puissances extérieures supérieures

Rien n'empêche par contre de continuer, *c'àd.* de regarder l'espace engendré par les triplets  $\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k$  en identifiant les permutations de  $(i, j, k)$  à un signe près (+ pour une permutation paire

et – pour une permutation impaire). Ceci permet de définir deux produits, le premier ayant pour entrée  $n - 1$  éléments de  $\mathbb{R}^n$  et pour sortie un élément de  $\mathbb{R}^n$  (i.e.  $(\mathbb{R}^n)^{(n-1)} \rightarrow \mathbb{R}^n$ ). Soit  $j \in \overline{n-1}$  et  $\mathbf{x}^{(j)} \in \mathbb{R}^n$ , il est défini par :

$$\mathbf{x}^{(1)} \wedge \mathbf{x}^{(2)} \wedge \dots \wedge \mathbf{x}^{(n-1)} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n-1)} & x_2^{(n-1)} & \dots & x_n^{(n-1)} \end{vmatrix},$$

et a pour propriété que,  $\forall j \in \overline{n-1}$ ,  $\mathbf{x}^{(j)} \cdot (\mathbf{x}^{(1)} \wedge \mathbf{x}^{(2)} \wedge \dots \wedge \mathbf{x}^{(n-1)}) = 0$ , c'ad. que son résultat est orthogonal à ses entrées. De manière amusante pour  $n = 2$ , le précédent produit est en fait une application  $J : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  ( $J(x_1, x_2) = (-x_2, x_1)$ ) qui est une anti-involution (i.e.  $\forall x \in \mathbb{R}^2$ ,  $J(J(x)) = -x$ ).

Le second (le produit mixte) prend quant à lui  $n$  éléments de  $\mathbb{R}^n$  pour leur associer un scalaire : soit  $j \in \overline{n}$  et soit  $\mathbf{x}^{(j)} \in \mathbb{R}^n$ ,

$$[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}] = \begin{vmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n)} & x_2^{(n)} & \dots & x_n^{(n)} \end{vmatrix},$$

## 1.B Espaces vectoriels abstraits et axiomes du choix

Soit  $K$  un corps (e.g.  $\mathbb{R}$  muni de  $+$  et  $\cdot$ ), dont les éléments seront appelés scalaires, et soit  $M$  un groupe (e.g.  $\mathbb{R}^n$  muni de la somme vectorielle  $+$ ). Si de plus il existe une multiplication par les scalaires (une fonction  $K \times M \rightarrow M$ ) de sorte que toutes ces opérations satisfont les propriétés EV1-8, alors  $M$  est dit un espace vectoriel sur  $K$ . Un produit scalaire est une opération  $M \times M \rightarrow \mathbb{R}$  qui satisfait les propriétés PS1-3, et ce produit scalaire induit toujours une norme définie par (1.3.1). L'inégalité de Cauchy-Schwarz 1.3.2 reste vraie dans ce contexte général. Elle s'utilise entre autres dans le traitement du signal (où  $M$  est alors un espace de fonctions de  $\mathbb{R} \rightarrow \mathbb{R}$ ).

Un exemple amusant est de prendre  $K = \mathbb{Q}$  et ensuite  $M = \mathbb{R}$ . En tant qu'espace vectoriel sur  $\mathbb{Q}$ ,  $\mathbb{R}$  ne peut pas admettre de base finie ; grâce à l'argument diagonal astucieux de Cantor, la base n'est même pas dénombrable (c'ad. même une suite infinie  $\mathbf{e}_1, \mathbf{e}_2, \dots$  ne suffit pas). Pour ceux qui admettent l'axiome du choix (un axiome mathématique qui est accepté par la plupart des mathématiciens), il est néanmoins possible de lui choisir une base (non-dénombrable), et cette base peut d'ailleurs être prise dans l'intervalle  $[0, 1]$ . Notons l'ensemble d'indices pour cette base  $I$ , i.e. la base est l'ensemble  $\{\mathbf{e}_i\}_{i \in I}$ . Ceci permet de faire quelques constructions amusantes.

La première est une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  telle que l'image de tout intervalle est dense dans  $\mathbb{R}$ . Sa construction est la suivante : dans notre base il y aura un seul élément dans  $\mathbb{Q}$ , disons  $\mathbf{e}_{i_0} = 1$ . Puisque c'est une base, tout réel  $r$  s'écrit de manière unique  $r = x_{i_0} \mathbf{e}_{i_0} + \mathbf{r}'$  où  $\mathbf{r}'$  est linéairement indépendant de  $\mathbf{e}_{i_0}$ . Alors, la définition de  $f$  est  $f(r) = x_{i_0} \mathbf{e}_{i_0} - \mathbf{r}'$ .

La seconde construction est l'ensemble de base lui-même. Il est possible de montrer que cet ensemble n'est pas mesurable, *c'ad.* qu'il est impossible de lui donner une taille qui soit consistante avec le fait que la taille de l'intervalle  $[a, b]$  soit  $b - a$ . Ceci n'est pas choquant en soi, mais c'est à partir de ces ensembles que le "paradoxe" de Banach-Tarski est fait. Ce "paradoxe"<sup>1</sup> dit qu'il est possible de découper une sphère en un nombre fini de sous-ensembles  $F_1, F_2, \dots, F_N$  puis de déplacer (par des rotations, symétries et translations) ces sous-ensembles pour reformer deux sphères identiques à la première. Ce "paradoxe" n'en est pas un ; c'est en fait un théorème en bonne et due forme, si l'axiome du choix est accepté.

Parmi les conséquences notoires de l'axiome du choix, citons le Lemme de Zorn, le théorème de compacité de Tychonoff, l'existence d'un bon ordre sur tous les ensembles, les ultrafiltres non-principaux, ... Il en existe des versions plus faibles (l'axiome du choix dénombrable) qui n'ont pas de conséquences désagréables et suffisent pour la plupart des mathématiques. Les mathématiques dites "constructivistes" se passent normalement de tout axiome du choix.

Parfois des implications subtiles communément admises nécessitent un axiome du choix : il y a deux définitions habituelles de la continuité pour les fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$ , celle dite de Cauchy (avec les  $\varepsilon$  et  $\delta$ ) et celle dite de Heine ( $\lim f(x_n) = f(\lim x_n)$ ). Si une fonction est Cauchy-continue en  $x$  elle est aussi Heine-continue. Si elle est Heine-continue sur tout un voisinage de  $x$  ou si  $x \in \mathbb{Q}$  elle est aussi Cauchy-continue en  $x$ . Cependant, en absence de l'axiome du choix, il n'est pas possible de montrer qu'une fonction Heine-continue en  $x \in \mathbb{R} \setminus \mathbb{Q}$  est Cauchy-continue en  $x$ ...

---

1. Les guillemets proviennent tout simplement du fait que ceci n'est absolument pas paradoxal : si l'axiome du choix est accepté, c'est en une conséquence et elle ne mène à aucune contradiction connue.



## Chapitre 2

# Applications, limites et continuité

*“La démonstration est la même, mutatis mutandis ...”*

-PROF.UNTEL

Soit  $S$  un sous-ensemble de  $\mathbb{R}^m$ . Une application  $\mathbf{F}$  de  $S$  dans  $\mathbb{R}^n$  est une règle qui associe à chaque élément  $\mathbf{x} \in S$  un élément  $\mathbf{F}(\mathbf{x}) \in \mathbb{R}^n$  (formellement, c'est un sous-ensemble de  $S \times \mathbb{R}^n$  tel que la projection sur le premier facteur est bijective). La représentation usuelle de  $\mathbf{F}$  se fait en coordonnées : pour  $i \in \overline{n}$ , soit  $F_i(\mathbf{x}) = \pi_i \circ \mathbf{F}(\mathbf{x})$ , appelée fonction coordonnée de  $\mathbf{F}$ , alors

$$(2.0.1) \quad \mathbf{F}(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ F_2(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{pmatrix}.$$

Ainsi l'étude des applications de  $\mathbb{R}^m$  dans  $\mathbb{R}^n$  se ramène à l'étude de  $n$  applications de  $\mathbb{R}^m$  dans  $\mathbb{R}$ . Pour certaines valeurs de  $m$  et  $n$ , d'autres termes que “applications” seront parfois préférés :

- pour  $n = 1$  le mot fonction,
- pour  $m = 1$  le terme courbe paramétrée,
- pour  $m = 2$  la locution surface paramétrée,
- pour  $m = n$  l'expression champ de vecteur.

Les courbes, surfaces et champ de vecteurs seront l'objet d'un traitement particulier dans les chapitres ultérieurs ; nous commencerons d'habitude par décrire ce qui se passe pour une fonction avant de passer au cas général d'une application. Tout ce qui sera dit pour une application s'applique évidemment à tous les autres cas. Cette séparation est faite car certaines notions géométriques sont particulières à certaines dimensions.

Pour ce qui est de la notation, les lettres majuscules grasses seront utilisées pour les applications, sauf lorsqu'il s'agit (à coup sûr) d'une

- fonction ( $n = 1$ ), dans quel cas une lettre majuscule en police normale sera utilisée, *e.g.*  $F, G, H, \dots$  ;
- fonction à valeur réelle ( $m = n = 1$ ) où la notation habituelle d'une lettre minuscule en police normale sera utilisée, *e.g.*  $f, g, h, \dots$  ;

## 2.1 - Représentations graphiques des fonctions

- courbe paramétrée ( $n = 1$ ), qui sera représentée par une lettre grecque minuscule, *e.g.*  $\gamma, \phi, \sigma$ , ... ;
- surface paramétrée ( $n = 2$ ), dont la notation sera une lettre grecque majuscule, *e.g.*  $\Phi, \Psi, \Xi$  ...

Finalement, lorsque  $m = 2$ , il est commun d'utiliser  $x$  et  $y$  plutôt que  $x_1$  et  $x_2$  comme nom des variables, et, lorsque  $m = 3$ ,  $x, y$  et  $z$  au lieu de  $x_1, x_2$  et  $x_3$ . Il va sans dire qu'il ne s'agit là que de notations, en particulier, elles sont probablement très différentes dans d'autres textes.

Deux notions qui se définissent avec une aise uniforme pour toutes les applications sont les limites et les questions de continuité. Le slogan pour passer des fonctions  $\mathbb{R} \rightarrow \mathbb{R}$  aux applications  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  est "remplacer les valeurs absolues  $|\dots|$  par des normes  $\|\dots\|$ ".

## 2.1 Représentations graphiques des fonctions

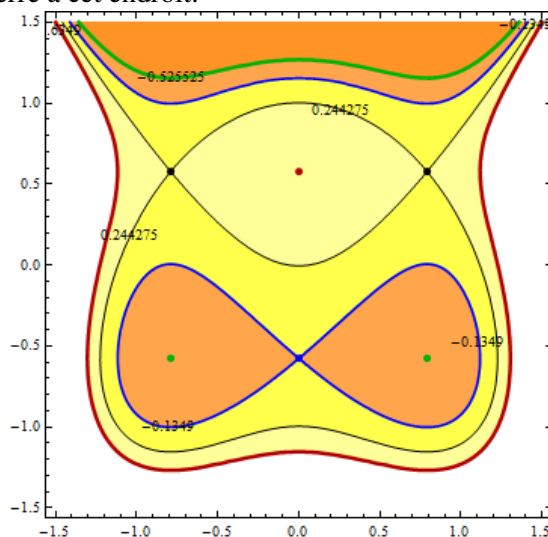
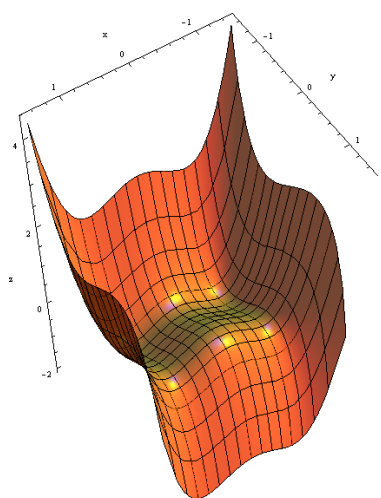
Comme le monde qui nous entoure n'est que 3 dimensionnel, les seules applications qu'il est aisé de représenter graphiquement (hormis bien sûr les fonctions  $\mathbb{R} \rightarrow \mathbb{R}$ ) sont les fonctions de  $\mathbb{R}^2$  ou  $\mathbb{R}^3$  dans  $\mathbb{R}$ , les courbes paramétrées de  $\mathbb{R}$  dans  $\mathbb{R}^2$  ou  $\mathbb{R}^3$  et les surfaces paramétrées de  $\mathbb{R}^2$  dans  $\mathbb{R}^3$ . Plus de détails sur les courbes et surfaces se trouvent dans le chapitre qui leur sera dédié.

Il y a essentiellement trois techniques communes : les graphes, les ensembles de niveau et les paramétrages.

La première ne peut réalistement être faite que pour une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$ , mais la notion de graphe existe en fait pour n'importe quelle application entre deux ensembles. "Ensemblistement", le graphe d'une application  $f : A \rightarrow B$  est  $\Gamma_f = \{(a, b) \in A \times B \mid a = f(b)\}$ . Pour une fonction  $F : S \rightarrow \mathbb{R}$  définie sur  $S \subset \mathbb{R}^2$ , ceci s'écrit

$$\Gamma_F = \{(x, y, z) \in S \times \mathbb{R} \mid z = F(x, y)\}$$

L'exemple typique serait de penser au domaine  $S$  de la fonction comme une parcelle de terrain et la valeur de  $F(x, y)$  comme l'altitude (par rapport au niveau de la mer) au point  $(x, y)$ . Le graphe d'une telle fonction aura l'air de la surface que décrit la terre à cet endroit.



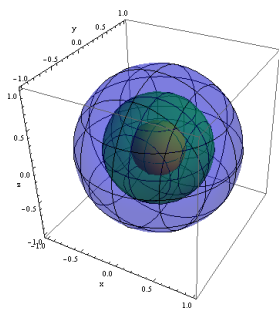
[À gauche la représentation de  $F(x, y)$  en graphe, à droite par les ensembles de niveau (les courbes, et parfois quelques points, dans le dessin).]

La seconde représentation est celle des ensembles de niveaux. Pour poursuivre dans l'idée de la fonction qui donne l'altitude d'une parcelle de terrain, les ensembles de niveaux seraient alors les courbes de niveaux de la carte topographique de cette parcelle. Mathématiquement, soit  $c \in \mathbb{R}$ , alors l'ensemble de niveau  $S_c$  de la fonction  $F : S \rightarrow \mathbb{R}$  définie sur  $S \subset \mathbb{R}^m$  est  $S_c = \{\mathbf{x} \in S \mid F(\mathbf{x}) = c\} = F^{-1}(c)$ .

Il est tout à fait possible de représenter les ensembles de niveaux d'applications de  $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tant que  $m \leq 3$ . Par contre, la dimension attendue (ici seulement en un sens intuitif) de l'ensemble de niveau de  $\mathbf{F}$  défini par  $F^{-1}(\mathbf{c})$  où  $\mathbf{c} \in \mathbb{R}^n$  est  $m - n$ . En particulier, lorsque  $n$  devient trop grand, la plupart des ensembles de niveaux de la fonction sont vides. Si  $m = n$  il ne s'agira ("génériquement") que de points. Deux cas "exotiques" qui se rencontrent relativement régulièrement sont ainsi

1.  $m = 3$  et  $n = 2$ . Les ensembles de niveaux auront la plupart du temps dimension 1, *c'àd.* ils formeront une famille de courbes paramétrées par deux réels : la courbe associée à  $(c_1, c_2)$  est définie par les points  $(x_1, x_2, x_3)$  satisfaisant  $\mathbf{F}(x_1, x_2, x_3) = (c_1, c_2)$ .
2.  $m = 4$  et  $n = 1$  : les ensembles de niveaux sont alors "génériquement" de dimension 3 ; chacun de ces ensembles est alors représenté comme une famille de surfaces. Pour décrire l'ensemble de niveau de  $c \in \mathbb{R}$ , il est alors commode de fixer (*c'àd.* de paramétrer par) une des coordonnées (par exemple  $x_4$ ). Il faut alors chercher la surface (dans  $\mathbb{R}^3$ ) donnée par  $F(x_1, x_2, x_3, x_4) = c$  et de regarder comment elle évolue lorsque  $x_4$  varie.

**Exemple 2.1.1.** Pour illustrer le premier cas "exotique", soit  $\mathbf{F} : U \rightarrow \mathbb{R}^2$  définie (sur  $U = \mathbb{R}^3 \setminus \{(0, 0)\} \times \mathbb{R}$ ) par  $\mathbf{F}(x, y, z) = (x^2 + y^2 + z^2, \arctg \frac{z}{\sqrt{x^2 + y^2}})$  alors si  $r \in \mathbb{R}_{>0}$  et  $\phi \in ]-\frac{\pi}{2}, \frac{\pi}{2}[$  alors  $F^{-1}(r, \phi)$  est un cercle centré en  $(0, 0, r \sin \phi)$  de rayon  $r \cos \phi$  et orienté de sorte que tout segment qui relie le centre à un de ses points est parallèle au plan  $xy$ .

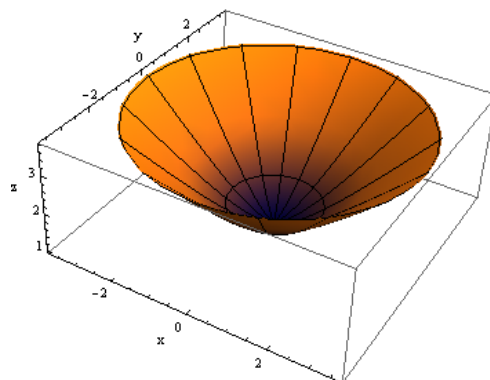
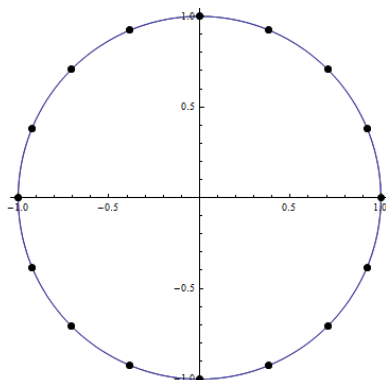


Pour le second, soit  $F : \mathbb{R}^4 \rightarrow \mathbb{R}$  définie par  $F(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - x_4^2$ . Pour  $c \in \mathbb{R}$ , l'ensemble de niveau  $\mathbf{F}(\mathbf{x}) = c$  se décrit bien comme suit : si  $x_4^2 > -c$  alors il s'agit d'une famille de sphères de rayon  $\sqrt{x_4^2 + c}$  (le dessin ci-contre représente par exemple  $c = 0$  et  $x_4 = \frac{1}{3}, \frac{2}{3}$  et 1). ♣

La dernière méthode de représentation est de ne regarder que l'image de la fonction. Elle porte le nom de paramétrage car il s'agit de décrire un ensemble non pas par une équation, mais comme ce qui est balayé en faisant varier des paramètres. La représentation graphique par paramétrage d'une application  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  où  $S \subset \mathbb{R}^m$  est alors  $\mathbf{G}(S) = \{\mathbf{y} \in \mathbb{R}^n \mid \exists \mathbf{x} \text{ tel que } \mathbf{F}(\mathbf{x}) = \mathbf{y}\} \subset \mathbb{R}^n$ . Pour faciliter la visualisation, il est alors fréquent de marquer certains points de l'image de manière particulière. Typiquement, quelque chose comme  $\mathbf{F}(\mathbb{Z} \cap S)$  si le domaine est contenu dans  $\mathbb{R}$ , ou encore,  $\mathbf{F}(\mathbb{Z} \times \mathbb{R} \cap S)$  et  $\mathbf{F}(\mathbb{R} \times \mathbb{Z} \cap S)$  si le domaine est contenu dans  $\mathbb{R}^2$ .

## 2.2 - Topologie dans $\mathbb{R}^n$

**Exemple 2.1.2.** La courbe paramétrée  $\mathbf{F} : \mathbb{R} \rightarrow \mathbb{R}^2$  définie par  $\mathbf{F}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$  décrit un cercle de rayon 1 dans  $\mathbb{R}^2$ .



[À gauche un paramétrage marqué (un point est fait à chaque  $\frac{\pi}{8}$  du cercle, à droite le paramétrage de l'hyperboloïde décrit ci-dessous.)]

Pour bien marquer comment est paramétrée la surface (un hyperboloïde) de

$$\mathbf{F}(r, \theta) = \begin{pmatrix} \operatorname{sh} r \cos \theta \\ \operatorname{sh} r \sin \theta \\ \operatorname{ch} r \end{pmatrix},$$

il est probablement judicieux de marquer les lignes  $\mathbf{F}(\mathbb{Z} \times \mathbb{R})$  et  $\mathbf{F}(\mathbb{R} \times \frac{\pi}{8}\mathbb{Z})$ . ♣

EXERCICE 8: Soit  $\bar{B}_3(\mathbf{0}) = \{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 \leq 9 \} \subset \mathbb{R}^2$  et  $F : B_3(\mathbf{0}) \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = \sqrt{9 - x^2 - y^2}$ . Représenter graphiquement les ensembles de niveau  $S_c = F^{-1}(c)$ , pour  $c = 0, \sqrt{5}, \sqrt{8}$  et 3.

## 2.2 Topologie dans $\mathbb{R}^n$

Afin de parler de limites, il faut introduire la notion qui remplacera l'intervalle autour d'un point  $\{y \in \mathbb{R} \mid |x - y| < \varepsilon\}$ , et comme le veut le slogan, son remplaçant sera la boule ouverte centrée en  $\mathbf{x} \in \mathbb{R}^n$  de rayon  $r$ . Pour ce faire, il faudra utiliser le fait qu'une norme induit (à son tour) une notion de distance entre deux points. Soient  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , alors la distance entre  $\mathbf{x}$  et  $\mathbf{y}$  est définie par la norme du vecteur qui les relie :

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \in \mathbb{R}_{\geq 0}.$$

Étant donné les propriétés précédentes, il est facile de vérifier que,  $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ ,

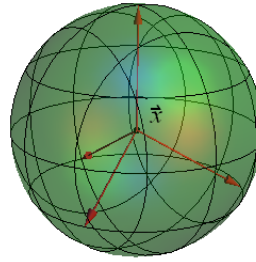
- Di1.  $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$  (symétrie)
- Di2.  $d(\mathbf{x}, \mathbf{y}) \geq 0$  (positivité)
- Di3.  $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$  (séparation)
- Di4.  $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$  (inégalité du triangle)



Ceci étant fait, les boules remplaceront la notion d'intervalle : la boule de rayon  $r$  en  $\mathbf{x}$  est

$$B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n \mid d(\mathbf{x}, \mathbf{y}) < r\}.$$

À noter que les boules ne contiennent pas leur bord (le terme boule ouverte est parfois aussi employé). Il est aussi bon de souligner que la boule de rayon 0 est l'ensemble vide (ce qui n'est pas



En topologie, les boules forment une base pour une topologie de  $\mathbb{R}^n$  (et permet aussi d'induire une topologie sur ses sous-ensembles). Ceci permet de faire une série de définitions à saveur topologique :

**Définition 2.2.1.** Soit  $S$  un sous-ensemble de  $\mathbb{R}^n$ .

1. Un point  $\mathbf{p} \in S$  est dit point isolé si  $\exists r \in \mathbb{R}_{>0}$  tel que  $B_r(\mathbf{p}) \cap S = \{\mathbf{p}\}$ .
2. Un point  $\mathbf{p} \in \mathbb{R}^n$  est dit point frontière (ou point de bord) de  $S$ , si  $\forall r \in \mathbb{R}_{>0}$ ,  $B_r(\mathbf{p}) \cap S \neq \emptyset$  et  $B_r(\mathbf{p}) \cap (\mathbb{R}^n \setminus S) \neq \emptyset$ .
3. Un point  $\mathbf{p} \in S$  est dit point d'intérieur de  $S$  si  $\exists r \in \mathbb{R}_{>0}$  tel que  $B_r(\mathbf{p}) \subset S$ .
4. Un point  $\mathbf{p} \in \mathbb{R}^n$  est dit point d'accumulation de  $S$  si  $\forall r \in \mathbb{R}_{>0}$ ,  $B_r(\mathbf{p}) \cap S \neq \emptyset$ .
5. Un point  $\mathbf{p} \in \mathbb{R}^n$  est un point limite de  $S$  si  $\mathbf{p}$  est un point d'accumulation mais n'est pas un point isolé.
6. Le bord de  $S$ , noté  $\partial S$ , est l'ensemble des points frontières de  $S$ .
7. L'intérieur de  $S$ , noté  $\overset{\circ}{S}$ , est l'ensemble des points d'intérieur de  $S$ .
8. L'adhérence de  $S$ , notée  $\bar{S}$ , est l'ensemble des points d'accumulation de  $S$ .
9.  $S$  est dit [un] ouvert si  $\overset{\circ}{S} = S$ .
10.  $S$  est dit [un] fermé si  $\bar{S} = S$ .
11. Un voisinage d'un point  $\mathbf{p}$  (resp. d'un ensemble  $S$ ) est un ouvert  $U$  tel que  $\mathbf{p} \in U$  (resp.  $S \subset U$ ). L'ensemble des voisinages de  $\mathbf{p}$  (resp.  $S$ ) est noté  $\text{Vois}\mathbf{p}$  (resp.  $\text{Vois}S$ ). ★

Attention : dans les définitions ci-dessus, un point d'accumulation de  $S$  (et par conséquent, un point limite de  $S$ ) n'appartient pas forcément à  $S$ . Par contre, dans le cours topologie, il est fort probable que la distinction ne soit pas faite entre point d'accumulation et point limite. D'ailleurs, cette différence ne devrait en général pas trop nous perturber, car les ensembles  $S$  seront le plus souvent des ouverts.

**Exemple 2.2.2.** L'ensemble  $S = \left\{ \left( \frac{1}{(n-1)/n} \right) \right\}_{n \in \mathbb{Z}_{\geq 1}} \subset \mathbb{R}^2$  qui n'est fait que de points isolés possède un seul point limite :  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ .

Par contre, l'ensemble  $\left\{ \begin{pmatrix} \sin n \\ \cos n \end{pmatrix} \right\}_{n \in \mathbb{Z}_{\geq 1}} \subset \mathbb{R}^2$  possède comme point limite tout point  $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$  tel que  $x^2 + y^2 = 1$  (pas évident à montrer). ♣

Quelques résultats de topologie seront maintenant énoncés ; les démonstrations ne seront pas faites car elles seront normalement abordées dans (et appartiennent plutôt à) un cours de topologie.



**Proposition 2.2.3**

Soit  $S \subset \mathbb{R}^n$ .

1.  $\mathring{S} = S \setminus \partial S$ .
2.  $S$  est ouvert si et seulement si  $S \cap \partial S = \emptyset$ .
3.  $\partial S \subset \bar{S}$ ,  $S \subset \bar{S}$  et  $\bar{S} = S \cup \partial S$ .
4.  $S$  est fermé si et seulement si  $\partial S \subset S$ .
5.  $S$  est fermé si et seulement si  $\mathbb{R}^n \setminus S$  est ouvert.
6.  $\partial S$  est un fermé et  $\mathring{S}$  est un ouvert.
7. Une union arbitraire d'ouverts est un ensemble ouvert.
8. Une intersection finie d'ouverts est un ouvert.

Un concept supplémentaire sera utilisé pour l'étude de fonctions au chapitre 4.

**Définition 2.2.4.** Un ensemble  $S$  est borné si  $\exists r \in \mathbb{R}_{>0}$  tel que  $S \subset B_r(\mathbf{0})$ .

Un ensemble  $K \subset \mathbb{R}^n$  est compact si et seulement si il est fermé et borné. ★

Le théorème suivant (qui est normalement la définition d'un compact en toute généralité) est mentionné par souci de complétude.

**Théorème 2.2.5**

$K$  est compact si et seulement si pour toute collection d'ouverts  $\{U_i\}_{i \in I}$  telle que  $\cup_{i \in I} U_i \supset K$ ,  $\exists n \in \mathbb{Z}_{\geq 0}$  et  $\exists \{i_k\}_{k \in \bar{n}}$  tel que  $\cup_{k \in \bar{n}} U_{i_k} \supset K$ .

Pour finir en douceur :

**Remarque 2.2.6.** Toute boule  $B_r(\mathbf{x}) \subset \mathbb{R}^m$  est contenue dans un ensemble “carré” (un produit d'intervalles, qui sera appelé plus tard un pavé). En effet, si  $\mathbf{y} \in B_r(\mathbf{x})$  alors  $|y_i - x_i| < r$ , ainsi  $\mathbf{y} \in \prod_{i \in \bar{m}} [x_i - r, x_i + r]$ . Inversement, soit  $r_i \in \mathbb{R}_{\geq 0}$  (où  $i \in \bar{m}$ ) et  $P = \prod_{i \in \bar{m}} [x_i - r_i, x_i + r_i]$ . Si  $\mathbf{y} \in P$  alors

$$\|\mathbf{y} - \mathbf{x}\| \leq \left( \sum_{i \in \bar{m}} r_i^2 \right)^{1/2} =: r. \text{ Ainsi } P \subset B_r(\mathbf{x}). \spadesuit$$

### 2.2.i Espace vectoriel vs espace topologique

Ce paragraphe a surtout pour but de souligner une particularité de  $\mathbb{R}^n$ , qui est celle d'être à la fois un espace topologique et un espace vectoriel (il y a même une certaine compatibilité entre les deux structures). Cependant, en parlant d'applications de  $\mathbb{R}^m$  dans  $\mathbb{R}^n$ , il est généralement préférable de ne penser qu'à l'aspect topologique de la chose. En effet, le domaine de l'application n'est en général pas tout  $\mathbb{R}^n$ , de la même façon que les fonctions des réels dans les réels ne sont pas, en général, définies sur tout  $\mathbb{R}$  et n'ont pas pour image tout  $\mathbb{R}$ .

Ainsi, lorsque la discussion aura lieu au niveau du domaine et de l'image de la fonction,  $\mathbb{R}^m$  ou ses sous-ensembles seront à proprement parlé un espace topologique. Par contre, comme l'idée fondatrice du cours est d'approcher les fonctions par des objets linéaires, il sera très souvent utile

d'utiliser  $\mathbb{R}^m$  comme un espace vectoriel (ou ses sous-espaces vectoriels). D'où la formule, légèrement alambiquée, de définir la distance entre deux *points* de  $\mathbb{R}^m$  (espace topologique) comme la norme du *vecteur* dans  $\mathbb{R}^m$  (espace vectoriel) qui les relie. D'où aussi l'insistance pour bien séparer un *point* (élément de l'espace topologique) d'un *vecteur* (élément de l'espace vectoriel), même si *a priori* il s'agit dans les deux cas d'un élément de  $\mathbb{R}^m$ .

## 2.3 Limites

Dire que la limite de  $\mathbf{F}(\mathbf{x})$  lorsque  $\mathbf{x}$  tend vers  $\mathbf{a}$  existe et vaut  $\mathbf{p}$  c'est, intuitivement, demander à ce que les valeurs de  $\mathbf{F}$  dans des voisinages de  $\mathbf{a}$  de plus en plus petits deviennent de plus en plus proches de  $\mathbf{p}$ . Ceci est formalisé par la définition de continuité (de Cauchy) suivante :

**Définition 2.3.1.** Soit  $S \subset \mathbb{R}^m$  et soit  $\mathbf{a} \in \mathbb{R}^m$  tel que  $\mathbf{a}$  est un point limite de  $S$ . L'application  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  admet la limite  $\mathbf{p} \in \mathbb{R}^n$  au point  $\mathbf{a}$ , notée  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{p}$ , si

$$\forall \varepsilon \in \mathbb{R}_{>0} \exists \delta \in \mathbb{R}_{>0} \text{ tel que } \forall \mathbf{x} \in (B_\delta(\mathbf{a}) \setminus \{\mathbf{a}\}) \cap S, \mathbf{F}(\mathbf{x}) \in B_\varepsilon(\mathbf{p}). \quad \star$$

En remplaçant les boules par leur définition et les distances par leur expression en norme, cette définition se résume à une incarnation du slogan :

$$\forall \varepsilon \in \mathbb{R}_{>0}, \exists \delta \in \mathbb{R}_{>0} \text{ tel que } \forall \mathbf{x} \in S \text{ satisfaisant } 0 < \|\mathbf{x} - \mathbf{a}\| < \delta, \text{ on ait } \|\mathbf{F}(\mathbf{x}) - \mathbf{p}\| < \varepsilon.$$

À noter que la définition de limite ci-dessus ne prend pas en compte la valeur au point  $\mathbf{a}$  en tant que tel, ni l'appartenance ou non de  $\mathbf{a}$  au domaine de définition de  $\mathbf{F}$ .

Cette section se termine sur quelques propriétés des limites ; leur démonstration est identique au cas des fonctions  $\mathbb{R} \rightarrow \mathbb{R}$  (à application du slogan près) d'où

EXERCICE 9: *Faire les démonstrations des énoncés de cette section. [Indice : aller voir la démonstration de l'énoncé correspondant pour les fonctions  $\mathbb{R} \rightarrow \mathbb{R}$ .]*

### Proposition 2.3.2

Soit  $m, n, \ell \in \mathbb{Z}_{\geq 1}$ , soit  $S \subset \mathbb{R}^m$  et  $S' \subset \mathbb{R}^n$ .

1. Soient  $\mathbf{F} : S \rightarrow S'$  et  $\mathbf{G} : S' \rightarrow \mathbb{R}^\ell$  deux applications telles que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{p}$  et  $\lim_{\mathbf{y} \rightarrow \mathbf{p}} \mathbf{G}(\mathbf{y}) = \mathbf{q}$ . Alors  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{G}(\mathbf{F}(\mathbf{x})) = \mathbf{q}$ .
2. Soit  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  une application telle que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{0}$  et  $G : S \rightarrow \mathbb{R}$  une fonction telle que  $|G(\mathbf{x})| \leq K$  pour un  $K \in \mathbb{R}$ . Alors  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x})G(\mathbf{x}) = \mathbf{0}$ .

Parmi les autres propriétés très utiles, notons le fait que la limite est déterminée par la limite sur chaque coordonnée. Un premier préliminaire est nécessaire, celui de parler de suites convergentes. Le concept de suite est toujours bien défini et utile dans  $\mathbb{R}^m$ . Il s'agit juste d'éléments de  $\mathbb{R}^m$  indicés par un entier positif, ou d'une application de  $\mathbb{Z}_{\geq 1} \subset \mathbb{R}$  dans  $\mathbb{R}^m$ .

**Définition 2.3.3.** Une suite  $\{\mathbf{x}_{(i)}\}_{i \in \mathbb{Z}_{\geq 1}}$  d'éléments de  $\mathbb{R}^m$  est dite convergente si l'application  $\mathbf{F} : S \rightarrow \mathbb{R}^m$ , où

$$S = \{c \in \mathbb{R} \mid c^{-1} \in \mathbb{Z}_{>0}\} = \{\frac{1}{k} \mid k \in \mathbb{Z}_{>0}\},$$

définie par  $\mathbf{F}(\frac{1}{k}) = \mathbf{x}_{(k)}$  admet une limite en  $0 \in \bar{S}$  (en fait 0 est le seul point limite de  $S$ ). La suite converge vers  $\mathbf{p}$  si  $\lim_{c \rightarrow 0} \mathbf{F}(c) = \mathbf{p}$ . La notation correspondante est  $\lim_{k \rightarrow \infty} \mathbf{F}(x_{(k)}) = \mathbf{p}$ . ★

Cette définition, un peu capillo-tractée, se ramène à une définition plus consensuelle :

#### Lemme 2.3.4

Une suite  $\{\mathbf{x}_{(i)}\}_{i \in \mathbb{Z}_{\geq 1}}$  converge vers  $\mathbf{p}$  si et seulement si

$$\forall \varepsilon \in \mathbb{R}_{>0} \exists N \in \mathbb{Z}_{\geq 1} \text{ tel que } \forall i \geq N, \|\mathbf{x}_{(i)} - \mathbf{p}\| < \varepsilon.$$

ou encore :

#### Lemme 2.3.5

Une suite  $\{\mathbf{x}_{(i)}\}_{i \in \mathbb{Z}_{\geq 1}}$  converge vers  $\mathbf{p}$  si et seulement si l'ensemble  $S = \{\mathbf{x}_{(i)}\}_{i \in \mathbb{Z}_{\geq 1}}$  est borné et  $\mathbf{p}$  est l'unique point limite de  $S$ .

Au passage, il vaut mieux mentionner que si  $K \subset \mathbb{R}^m$  est compact alors toute suite  $\{\mathbf{x}_{(n)}\}_{n \in \mathbb{Z}_{\geq 1}}$  admet une sous-suite  $\{\mathbf{x}_{(n_k)}\}_{k \in \mathbb{Z}_{\geq 1}}$  convergente. Il faut vérifier d'autres énoncés naturels :

#### Lemme 2.3.6

Soit  $i \in \bar{n}$  et  $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction coordonnée, et  $\{\mathbf{x}_{(i)}\}_{i \in \mathbb{Z}_{\geq 1}}$  une suite convergeant vers  $\mathbf{p}$ . Alors la suite  $\pi_k(\mathbf{x}_{(i)})$  converge vers  $\pi_i(\mathbf{p}) =: p_i$ .

Une petite application de l'inégalité triangulaire donne ensuite :

#### Proposition 2.3.7

Soit  $S \subset \mathbb{R}^m$ , soit  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  une application et soit  $F_i = \pi_i \circ \mathbf{F}$  alors la limite  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x})$  existe et vaut  $\mathbf{p} = (p_1, p_2, \dots, p_n)^T$  si et seulement si, pour tout  $i \in \bar{n}$ ,  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} F_i(\mathbf{x})$  existe et vaut  $p_i$ .

En réalité, les propriétés précédentes permettent de déduire à très peu de frais, les propriétés suivantes (et beaucoup d'autres).

#### Corollaire 2.3.8

Soient  $\mathbf{F}, \mathbf{G} : S \rightarrow \mathbb{R}^n$  deux applications telles que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{p}$  et  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{G}(\mathbf{x}) = \mathbf{q}$ . Alors  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} (\mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})) = \mathbf{p} + \mathbf{q}$ .

Soient  $\mathbf{F}, \mathbf{G} : S \rightarrow \mathbb{R}^n$  deux applications telles que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{p}$  et  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{G}(\mathbf{x}) = \mathbf{q}$ . Alors  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} (\mathbf{F}(\mathbf{x}) \cdot \mathbf{G}(\mathbf{x})) = \mathbf{p} \cdot \mathbf{q}$ .

Soit  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  une application telle que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{p}$  et  $G : S \rightarrow \mathbb{R}$  une fonction telle que  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} G(\mathbf{x}) = q \neq 0$ . Alors  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\mathbf{F}(\mathbf{x})}{G(\mathbf{x})} = \frac{1}{q} \mathbf{p}$ .

Finalement, voici un critère très utile pour montrer qu'il n'y a pas continuité :

#### Proposition 2.3.9

Étant donné deux suites  $\{\mathbf{x}_{(k)}\}_{k \in \mathbb{Z}_{\geq 0}}$  et  $\{\mathbf{y}_{(k)}\}_{k \in \mathbb{Z}_{\geq 0}}$ , si

$$\lim_{k \rightarrow \infty} \mathbf{x}_{(k)} = \mathbf{p} = \lim_{k \rightarrow \infty} \mathbf{y}_{(k)} \text{ et } \lim_{k \rightarrow \infty} \mathbf{F}(\mathbf{x}_{(k)}) \neq \lim_{k \rightarrow \infty} \mathbf{F}(\mathbf{y}_{(k)}),$$

alors la limite de  $\mathbf{F}$  en  $\mathbf{p}$  n'existe pas.

Par contraposition, si la limite de  $\mathbf{F}$  en  $\mathbf{p}$  existe,  $\lim_{k \rightarrow \infty} \mathbf{F}(\mathbf{x}_{(k)}) = \lim_{k \rightarrow \infty} \mathbf{F}(\mathbf{y}_{(k)})$ . En particulier, si une limite existe, il est complètement inutile de regarder les suites pour en montrer l'existence : il est impossible d'essayer toutes les suites possibles...

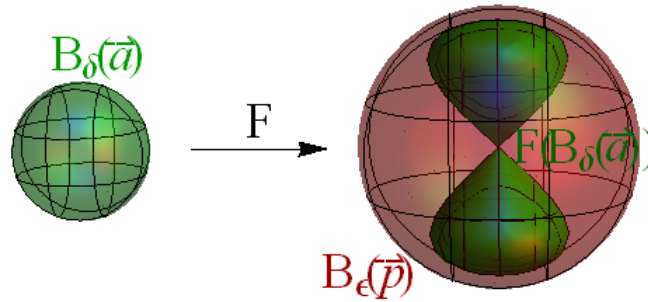
**Exemple 2.3.10.** Soit  $F(x, y) = \frac{xy}{x^2+y^2}$  définie pour tout  $\begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}$ . Cette fonction ne possède pas de limite en  $\mathbf{0}$ . En effet,  $F(\frac{1}{k}, \frac{1}{k}) = \frac{1}{2}$  et  $F(\frac{1}{k}, -\frac{1}{k}) = -\frac{1}{2}$ . ♣

## 2.4 Continuité

**Définition 2.4.1.** Soit  $S \subset \mathbb{R}^m$  et soit  $\mathbf{a} \in S$ . Une application  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  est dite continue en  $\mathbf{a} \in S$  un point limite si et seulement si

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{a}).$$

Si  $\mathbf{a}$  est un point isolé, alors la fonction est continue en  $\mathbf{a}$ . Elle est dite continue sur  $S' \subset S$  si elle est continue en tout point de  $S'$ . Si la limite existe mais ne vaut pas  $\mathbf{F}(\mathbf{a})$  alors l'application a une discontinuité réparable en  $\mathbf{a}$ . Si la limite n'existe pas, la locution discontinuité irréparable est alors employée. ★



Le nom de discontinuité réparable provient de ce que l'application peut alors être modifiée (en changeant sa valeur en  $\mathbf{a}$ ) pour devenir une application continue en ce point. Les résultats principaux sur la continuité sont tous corollaires des propriétés des limites.

### Corollaire 2.4.2

Soit  $S \subset \mathbb{R}^n$ , soit  $S' \subset \mathbb{R}^m$ , soit  $\mathbf{F} : S \rightarrow S'$  continue en  $\mathbf{a} \in S$  et  $\mathbf{G} : S' \rightarrow \mathbb{R}^\ell$  continue en  $\mathbf{F}(\mathbf{a})$ . Alors  $\mathbf{G} \circ \mathbf{F} : S \rightarrow \mathbb{R}^\ell$  est continue en  $\mathbf{a}$ . En particulier, la somme et le produit de fonctions continues sont continues.

**Exemple 2.4.3.** L'application

$$\mathbf{F}(x, y, z) = \begin{pmatrix} \cos(x+y) - e^z \\ x^3 + \sin(2xz + y^2) \\ \log(1 + x^2 + y^2) \end{pmatrix}$$

est continue. En effet, chacune de ses fonctions coordonnées est la somme/produit/composition de fonctions continues. La troisième coordonnée est légèrement plus délicate à traiter (mais comme  $x^2 + y^2 \geq 0$ , et que le log est continu sur  $\mathbb{R}_{\geq 1}$ , il n'y a pas de problèmes). ♣

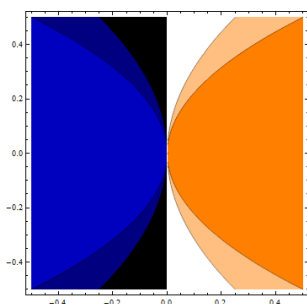


La liberté qu'offre la dimension supérieure joue parfois des tours au calculateur trop cavalier. En effet, pour une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$ , il suffit de vérifier que la limite à gauche et à droite concorde pour savoir que la fonction a, dans le pire des cas, une discontinuité réparable. Ceci est malheureusement faux en dimension supérieure.

**Exemple 2.4.4.** Soit  $F(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = \begin{cases} y^2/x & \text{si } |x| > 0, \\ 0 & \text{si } |x| = 0. \end{cases}$$

Le long d'une droite passant par l'origine cette fonction semble continue : si  $\{t_k\}_{k \in \mathbb{Z}_{\geq 0}}$  est une suite de réels de limite 0, alors pour tout  $c_1, c_2 \in \mathbb{R}$ ,  $\lim_{k \rightarrow \infty} F(c_1 t_k, c_2 t_k) = 0$ . Cependant, ce n'est pas le cas :  $\lim_{k \rightarrow \infty} F(t_k^2, t_k) = 1$  et  $\lim_{k \rightarrow \infty} F(2t_k^2, t_k) = 1/2$ .



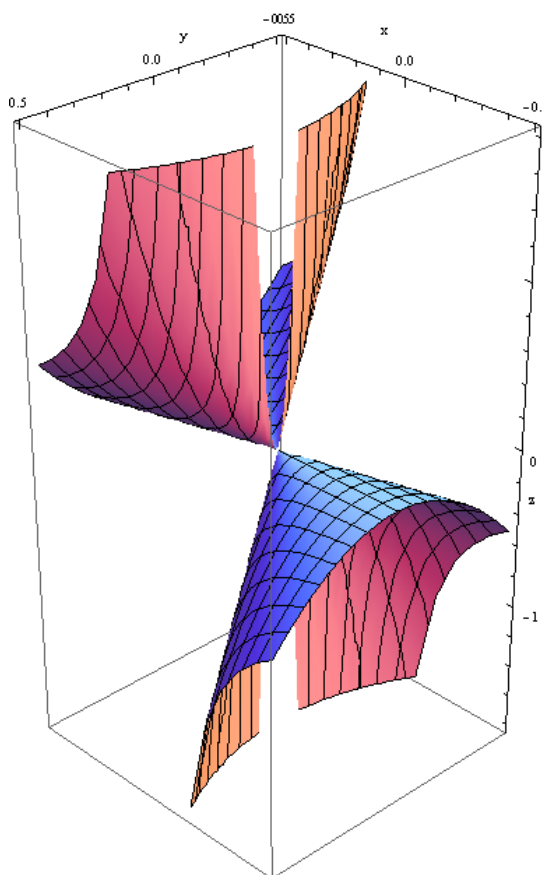
De fait, dans la représentation du graphe ci-contre et en surface de niveau ci-dessus, il est clair que la fonction n'est pas continue en ce point. ♣

Il arrive fréquemment que l'expression d'une fonction par une formule ne couvre pas clairement tous les cas. Ou alors que la fonction soit définie sur un ouvert et qu'il est souhaitable de définir sa valeur au bord de cet ouvert. De fait, dans la définition 2.3.1, seuls les points  $\mathbf{a} \in S$  qui sont des points limites sont considérés (alors que les points limites peuvent ne pas appartenir à  $S$ ). La terminologie habituelle dans ce cas est de parler d'extension par continuité :

**Exemple 2.4.6.** Soit  $F : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = \frac{x^3}{x^2 + y^2}$ . Comme  $F(0, y) = 0$ , il est naturel de tenter de l'étendre par 0 en  $\mathbf{0}$ . Tentons de montrer que  $\lim_{\mathbf{x} \rightarrow \mathbf{0}} F(x, y) = 0$ . D'une part,

**Définition 2.4.5.** Soit  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  et soit  $\mathbf{a} \notin S$ . Si  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{p}$ , alors l'extension de  $\mathbf{F}$  par continuité en  $\mathbf{a}$  est l'application qui associe à  $\mathbf{x} \in S \cup \{\mathbf{a}\}$  la valeur  $\mathbf{F}(\mathbf{x})$  si  $\mathbf{x} \in S$  et  $\mathbf{p}$  si  $\mathbf{x} = \mathbf{a}$ . ★

Il est alors possible d'étendre considérablement le domaine de définition de  $\mathbf{F}$  par continuité. En d'autres termes, aux points  $\mathbf{p}$  qui sont des points limites de  $S$  où  $\mathbf{F}$  n'est pas définie, la valeur  $\mathbf{F}(\mathbf{p}) = \lim_{\mathbf{x} \rightarrow \mathbf{p}} \mathbf{F}(\mathbf{x})$  est donnée à  $\mathbf{F}$  (si la limite existe !). Il ne s'agit que de donner une valeur (la seule possible lorsque la limite existe) de sorte que la fonction est continue en ce point.



$\|\mathbf{x} - \mathbf{0}\| < \delta \Leftrightarrow x^2 + y^2 < \delta^2$  et de l'autre  $x^2 \leq x^2 + y^2 \Rightarrow |x|^3 \leq (x^2 + y^2)^{3/2}$ , ainsi

$$|F(x, y)| = \left| \frac{x^3}{x^2 + y^2} \right| \leq \left| \frac{(x^2 + y^2)^{3/2}}{x^2 + y^2} \right| = (x^2 + y^2)^{1/2} < \delta.$$

Il suffit, pour tout  $\varepsilon$  donné, de poser  $\delta = \varepsilon$  dans la définition 2.3.1 pour voir que la limite existe bel et bien. Poser  $F(0, 0) = 0$  est une extension de  $F$  en  $\mathbf{0}$  par continuité. ♣

Finalement, voici une autre caractérisation de la continuité :

### **Théorème 2.4.7**

Soit  $F : S \rightarrow \mathbb{R}^n$  une application.  $F$  est continue si et seulement si  $\forall U \subset \mathbb{R}^n$  ouvert, l'ensemble  $F^{-1}(U)$  est un ouvert.

EXERCICE 10: Montrer le théorème précédent.

Les théorèmes 2.2.5 et 2.4.7 se conjuguent pour donner le corollaire suivant :

### **Corollaire 2.4.8**

Si  $K$  est compact et  $F$  continue alors  $F(K)$  est compact.

Le dernier corollaire se déduit du précédent et du théorème 2.2.5. S'il est besoin de le préciser un maximum absolu (resp. minimum absolu) d'une fonction  $F : S \rightarrow \mathbb{R}$  est un point  $\mathbf{p}$  de  $S$  tel que  $\forall \mathbf{x} \in S, F(\mathbf{x}) \leq F(\mathbf{p})$  (resp.  $F(\mathbf{x}) \geq F(\mathbf{p})$ ).

### **Corollaire 2.4.9**

Soit  $S \subset \mathbb{R}^n$  compact. Toute fonction  $F : S \rightarrow \mathbb{R}$  admet un maximum absolu et un minimum absolu.

EXERCICE 11: En utilisant la définition de la continuité de Cauchy (avec les  $\varepsilon$  et  $\delta$ ), montrer que la fonction  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  définie par :  $F(x, y) = 2xy + x^2$  est continue en  $(1, 1)$ . [Indice :  $2xy + x^2 - 3 = 2(x - 1)y + (x - 1)^2 + 2(y - 1) + 2(x - 1)$ ].

EXERCICE 12: Soit  $F : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = \frac{x^2 - x^3 + y^2}{x^2 + y^2}$ . Est-il possible d'étendre  $F$  par continuité en  $\mathbf{0}$  ?

[Indice : démontrer d'abord  $|x|^3 \leq (x^2 + y^2)^{3/2}$ .]

EXERCICE 13: Soit  $F : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = \frac{x^4 + y^4 - x^5}{x^4 + y^4}$ . Est-il possible d'étendre  $F$  par continuité en  $\mathbf{0}$  ?

[Indice : démontrer d'abord  $|x|^5 \leq (x^4 + y^4)^{5/4}$  et  $(x^4 + y^4)^{1/4} \leq (x^2 + y^2)^{1/2}$ .]

EXERCICE 14: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = \begin{cases} \frac{7x^4y^4 + 2xy^7 - \pi x^5y^3}{x^8 + y^8} & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}, \\ 0 & \text{si } x = y = 0. \end{cases}$$

Montrer que la fonction  $F$  n'est pas continue en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Cette discontinuité est-elle réparable ?

EXERCICE 15: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = \begin{cases} \log\left(1 + \frac{|xy|}{x^2 + y^2}\right) & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}, \\ 0 & \text{si } x = y = 0. \end{cases}$$

Montrer que  $\lim_{x \rightarrow 0} F(x, y) = 0$  et  $\lim_{y \rightarrow 0} F(x, y) = 0$ . Est-il possible d'en déduire que  $\lim_{(x, y) \rightarrow \mathbf{0}} F(x, y) = 0$  ?

EXERCICE 16: Pour  $m \in \mathbb{Z}_{\geq 1}$  et  $i \in \overline{m}$ , la projection canonique  $\pi_i$  de  $\mathbb{R}^m$  dans  $\mathbb{R}$  est définie par,  $\pi_i(\mathbf{x}) = x_i$ . Montrer que cette fonction est continue.

EXERCICE 17: Soit  $\mathbf{F}(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  l'application définie par

$$\mathbf{F}(x, y) = \begin{cases} \left( \frac{x^4+y^4}{x^2+y^2}, e^x \sin y \right)^T & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}; \\ \mathbf{0} & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0}. \end{cases}$$

Montrer qu'elle est continue sur  $\mathbb{R}^2$ .

[Indice : Regarder séparément chaque fonction coordonnée  $F_1$  et  $F_2$ .]

EXERCICE 18: Soit  $F(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = \begin{cases} \ln \left( 1 + \frac{x^4 y^4 + 2x^2 y^6 + 7x^6 y^2}{x^6 + y^6} \right) & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0 & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{cases}$$

a. Écrire la fonction  $G(x, y)$  tel que  $F(x, y) = \ln(1 + G(x, y))$ .

b. Montrer que  $F$  est continue sur  $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ .

c. Montrer que  $F$  est continue en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .

EXERCICE 19: Un sous-ensemble  $S$  de  $\mathbb{R}^m$  est dit fermé, si et seulement si  $\forall x \notin S, \exists r > 0$  tel que  $B_r(\mathbf{x}) \cap S = \emptyset$ . Montrer que l'ensemble  $\mathcal{T}$  des fermés de  $\mathbb{R}^m$  vérifie les propriétés suivantes :

Fe1. L'ensemble  $\emptyset$  et  $\mathbb{R}^n$  sont fermés ;

Fe2. Toute intersection de fermés est fermée ;

Fe3. L'union de deux fermés est fermée.

[Il est aussi possible de parler de la topologie en regardant ses ensembles fermés, e.g. dans la topologie de Zariski.]

EXERCICE 20: Montrer qu'un intervalle fermé de  $\mathbb{R}$  est un fermé de  $\mathbb{R}$  au sens ci-dessus, puis trouver une union infinie d'intervalles fermés qui n'est pas un ensemble fermé de  $\mathbb{R}$  (au sens ci-dessus).



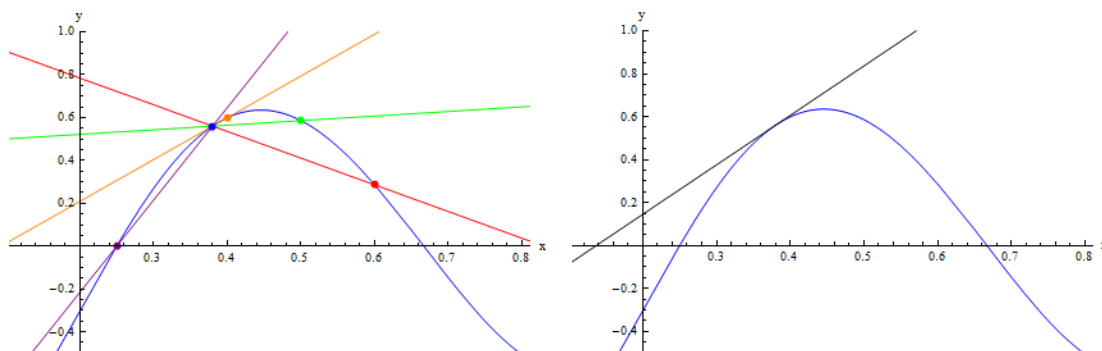
## Chapitre 3

# Dérivées et dérivabilité

“Dériver, c’est du gâteau.”

-TOTO

Pour les fonctions  $\mathbb{R} \rightarrow \mathbb{R}$ , la dérivée peut être perçue comme la variation instantanée. En effet, la variation moyenne d’une fonction entre deux réels  $a$  et  $b$  se définit raisonnablement par  $\frac{f(b)-f(a)}{b-a}$ . La convention de signe est heureuse dans la mesure où si la fonction est croissante (resp. décroissante), le chiffre obtenu est positif (resp. négatif) quand  $a < b$  tout comme quand  $a > b$ . Cette convention est d’autant plus utile car la droite  $y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$  est une approximation grossière de la fonction sur cet intervalle (encore une fois indépendamment du signe de  $b-a$ ). En fait pour la plupart des fonctions lisses (l’emploi d’un vocabulaire flou est volontaire), plus l’intervalle est petit plus l’approximation est correcte. L’expression moderne (*sic*) de la variation instantanée en  $x$  s’obtient en posant  $a = x$  et  $b = x + h$ , puis en prenant  $h$  de plus en plus petit, *i.e.*  $\lim_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h}$ . (La convention plus haut fait que le signe  $h = b-a$  n’est pas important.) Voici quelques sécantes (les droites de la forme  $y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ ) dessinées sur un graphe (à gauche,  $a$  est fixé et  $b$  varie) et la tangente (même équation mais le rapport est remplacé par  $f'(a)$ , à droite).



Pour une fonction  $F : \mathbb{R}^m \rightarrow \mathbb{R}$  à plusieurs variables ( $m > 1$ ) tenter de cerner la multitude de variations possibles autour d’un point est tâche impossible. Alors que dans  $\mathbb{R}$  il n’y a que deux façons de s’approcher continûment d’un point  $x$ , il y a en  $\mathbb{R}^m$  une infinité étourdissante pour  $\mathbf{x} \in \mathbb{R}^m$ . Par infinité étourdissante, le lecteur avisé comprendra qu’il y en a au moins autant que de fonctions continues de  $\mathbb{R} \rightarrow \mathbb{R}$ . De fait, en supposant  $m = 2$ , pour chaque fonction continue  $f : \mathbb{R} \rightarrow \mathbb{R}$ , on



$D_i \mathbf{F}(\mathbf{y})$  et  $D_i \mathbf{G}(\mathbf{y})$  existent. Soit  $\mathbf{H} : U \rightarrow \mathbb{R}^n$  l'application définie par

1.  $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})$ , alors  $D_i \mathbf{H}(\mathbf{y}) = D_i \mathbf{F}(\mathbf{y}) + D_i \mathbf{G}(\mathbf{y})$  ;
2.  $\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \cdot \mathbf{G}(\mathbf{x})$ , alors  $D_i \mathbf{H}(\mathbf{y}) = D_i \mathbf{F}(\mathbf{y}) \cdot \mathbf{G}(\mathbf{y}) + \mathbf{F}(\mathbf{y}) \cdot D_i \mathbf{G}(\mathbf{y})$ .

Le cas des fonctions composées est généralement plus subtil et sera traité dans une section ultérieure.

D'ailleurs, il arrive fréquemment que, plutôt que d'appeler les variables  $x_i$ , celles-ci soient représentées par d'autres symboles (comme  $x, y$  et  $z$ ). Si, par exemple,  $\mathbf{G}(\rho, \theta, \phi)$  est une application de  $U \subset \mathbb{R}^3$  dans  $\mathbb{R}^3$ , les notations changent pour rester cohérentes avec ces nouveaux noms :  $D_2 \mathbf{G}(\rho, \theta, \phi) = \frac{\partial}{\partial \theta} \mathbf{G}(\rho, \theta, \phi)$ .

**Exemple 3.1.4.** Soit  $\mathcal{U} : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $\mathcal{U}(\star, \varnothing) = \ln(1 + \star^2 + \varnothing^2)$ , alors

$$D_1 \mathcal{U}(\star, \varnothing) = \frac{\partial}{\partial \star} \mathcal{U}(\star, \varnothing) = \frac{2\star}{1 + \star^2 + \varnothing^2} \quad \text{et} \quad D_2 \mathcal{U}(\star, \varnothing) = \frac{\partial}{\partial \varnothing} \mathcal{U}(\star, \varnothing) = \frac{2\varnothing}{1 + \star^2 + \varnothing^2}. \quad \clubsuit$$

**Remarque 3.1.5.** Notons qu'il pourrait y avoir ambiguïté dans la notation  $\frac{\partial}{\partial x_i}$  lorsque des fonctions composées sont prises en considération. En effet, prenons la fonction  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par  $F(x, y) = x^2 y$ . Ses dérivées partielles sont définies en tout point de  $\mathbb{R}^2$  :  $(D_1 F)(x, y) = 2xy$  et  $(D_2 F)(x, y) = x^2$ . Que veut dire  $D_1 F(x^2, xy)$  ou  $\frac{\partial}{\partial x} F(x^2, xy)$  ? s'agit-il de  $D_1 F$  évaluée en  $(x^2, xy)$  ou de la fonction  $G(x, y) = F(x^2, xy) = x^5 y$  dérivée selon sa première variable ? La distinction suivante sera faite :

$$\begin{aligned} D_1 F(x^2, xy) &= (D_1 F)(x^2, xy) = \left( \frac{\partial F}{\partial x} \right)(x^2, xy) = (\partial_x F)(x^2, xy) = 2x^3 y, \\ \frac{\partial}{\partial x} F(x^2, xy) &= \frac{\partial}{\partial x} (F(x^2, xy)) = \partial_x (F(x^2, xy)) = 5x^4 y. \end{aligned}$$

L'emploi des parenthèses est de mise pour éradiquer le doute. ♠

Les dérivées partielles comportent une lacune que la dérivée d'une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  ne possède pas. Effectivement, lorsque la fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  a une dérivée en  $x \in \mathbb{R}$ , alors elle est continue en  $x$ . Malheureusement, en dimension supérieure, l'existence de dérivées partielles ne suffit pas pour avoir la continuité.

EXERCICE 21: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = \begin{cases} 0 & \text{si } x = y = 0, \\ \frac{xy}{x^2 + y^2} & \text{sinon.} \end{cases}$$

Montrer que en  $(0, 0)$  les deux dérivées partielles de  $F$  existent, mais  $F$  n'est pas continue en ce point (et que la discontinuité n'est pas réparable).

Un autre exemple est la fonction de l'exemple 2.4.4 : la fonction restreinte à n'importe quelle droite passant par l'origine est dérivable, et pourtant, elle n'est même pas continue !

EXERCICE 22: Trouver les dérivées partielles de

1.  $F(x, y) = \ln(1 + x^4 + y^4)$  ;
2.  $\mathbf{G}(\rho, \theta, \phi) = (\rho \sin \theta, \theta - \rho \cos \theta, \arctg \phi)$  ;

### 3.2 - Application linéaire et dérivabilité

$$3. H(x_1, x_2, x_3, x_4) = \frac{x_1 - x_2^2}{x_3^2 + x_4^4} \text{ si } x_3^2 + x_4^2 \neq 0;$$

$$4. T(x, y, z) = e^{x^2 y^3 z^4};$$

$$5. \Xi(x, y, z) = \ln(2 + \cos(xyz))$$

EXERCICE 23: Soit  $T : (\mathbb{R} \setminus \{0\}) \times \mathbb{R} \rightarrow \mathbb{R}$  la fonction définie par

$$T(x, y) = \begin{cases} \arctg \frac{y}{x} & \text{si } x > 0, \\ \pi + \arctg \frac{y}{x} & \text{si } x < 0, \end{cases}$$

Calculer  $D_1 T$  et  $D_2 T$ . Ces deux fonctions ne sont a priori pas définies si  $x = 0$ . Peuvent-elles s'étendre par continuité aux points  $\begin{pmatrix} 0 \\ y_0 \end{pmatrix}$ ? [Attention : votre réponse doit dépendre de  $y_0$ .]

Montrer que  $T$  ne peut pas s'étendre par continuité en  $\begin{pmatrix} 0 \\ y_0 \end{pmatrix}$  si  $y_0 \leq 0$ .

Avant de poursuivre, voici un rappel important de l'analyse en une variable.

#### **Théorème 3.1.6 (Théorème de la valeur Intermédiaire)**

Soit  $f : [a, b] \rightarrow \mathbb{R}$  une fonction continue et dérivable sur  $[a, b]$ . Alors pour tout  $h \in [0, b - a]$ ,  $\exists s \in [0, 1]$  tel que  $f(a + h) - f(a) = hf'(a + sh)$ .

## 3.2 Application linéaire et dérivabilité

L'idée d'approcher une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  par une droite est en fait la notion la plus porteuse pour parler de dérivabilité : une fonction  $f$  est dérivable en  $x$  si elle s'approche "bien" par une droite en ce point. Il faut tout de même deviner le bon concept qui remplace une droite ; comme le paragraphe précédent le montre, il ne suffit pas de regarder toutes les droites passant par un point. Il s'agira tout simplement d'un espace linéaire (plan, hyperplan...) incarné ici par une application linéaire, notion dont la définition est maintenant rappelée.

**Définition 3.2.1.** Une application  $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$  est dite linéaire si  $\forall a, b \in \mathbb{R}$  et  $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m$

$$L(a\mathbf{x} + b\mathbf{y}) = aL(\mathbf{x}) + bL(\mathbf{y}). \quad \star$$

Ainsi, il sera d'autant plus important de maintenir la distinction entre point (qui est "mangé" par une application quelconque) et vecteur (qui ne peut être "mangé" que par une application linéaire). Le théorème suivant permet de relier cette définition abstraite à une construction plus familière : les matrices.

#### **Théorème 3.2.2**

Une application  $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$  est linéaire si et seulement si il existe une matrice  $A$  de taille  $n \times m$  telle que  $L(\mathbf{x}) = A\mathbf{x}$ , ou, en coordonnées,

$$\begin{aligned} L_1(\mathbf{x}) &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ L_2(\mathbf{x}) &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ &\vdots \\ L_n(\mathbf{x}) &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m \end{aligned}$$

**DÉMONSTRATION IDÉE DE LA PREUVE :** C'est un résultat crucial (et élémentaire) de l'algèbre linéaire qu'une matrice est une application linéaire. Étant donné une application linéaire, il suffit de construire sa matrice en posant que  $\mathbf{L}(\mathbf{e}_i)$  est la  $i^{\text{ème}}$  ligne de cette matrice, puis d'utiliser (1.1.1) et la linéarité pour conclure que la matrice donne bien la même application que  $\mathbf{L}$ . ■

**Exemple 3.2.3.** Soit  $L : \mathbb{R}^m \rightarrow \mathbb{R}$  une application linéaire, alors sa matrice est une ligne  $A = (a_1, a_2, \dots, a_m)$ . En particulier,  $A^T$  est une colonne et donc s'identifie à un vecteur de  $\mathbb{R}^m$ ,  $\mathbf{a} = A^T$ . En outre, l'application peut s'exprimer par le biais du produit scalaire  $L(\mathbf{x}) = \mathbf{A}\mathbf{x} = \mathbf{a} \cdot \mathbf{x}$ . ♣

Dans les trois exercices suivants, il sera discuté des normes sur des matrices, un concept utile pour justifier la définition de dérivabilité. L'ensemble  $\mathcal{M}_{n \times m}$  des matrices (à coefficients dans  $\mathbb{R}$ ) possède la structure d'un espace vectoriel (sur  $\mathbb{R}$ ). En effet, soit  $A$  et  $B$  des matrices  $n \times m$ ,  $a_{ij}$  et  $b_{ij}$  leurs coefficients et  $\lambda$  un réel. Alors la multiplication par un scalaire est définie par le fait que le coefficient  $ij$  de  $\lambda A$  est  $\lambda a_{ij}$  et la somme en disant que le coefficient  $ij$  de  $A + B$  est  $a_{ij} + b_{ij}$ . À chaque matrice, il est *a priori* possible de lui associer deux réels :

$$\|A\|_{\rightarrow} = \sup_{\mathbf{v} \in \mathbb{R}^m \setminus \{0\}} \frac{\|A\mathbf{v}\|}{\|\mathbf{v}\|} \quad \text{et} \quad \|A\|_2 = \left( \sum_{i \in \overline{m}, j \in \overline{n}} |a_{ij}|^2 \right)^{1/2}.$$

EXERCICE 24: Montrer que  $\|\cdot\|_2$  est une norme. [Indice : identifier  $\mathcal{M}_{n \times m}$  à  $\mathbb{R}^{mn}$ .]

EXERCICE 25: Montrer que  $\|\cdot\|_{\rightarrow}$  est une norme. Si de plus  $A$  et  $B$  sont des matrices  $m \times m$ , montrer que  $\|AB\|_{\rightarrow} \leq \|A\|_{\rightarrow} \|B\|_{\rightarrow}$  (propriété dite de sous-multiplicativité).

EXERCICE 26: Montrer que  $\|A\|_{\rightarrow} \leq \|A\|_2$ . [Indice : utiliser Cauchy-Schwartz.]

La notion de dérivabilité sera que l'application s'approche bien par une application linéaire. Plus précisément,  $\mathbf{L}$  est un candidat pour une approximation linéaire de  $\mathbf{F}$  en  $\mathbf{x}$  si  $\mathbf{F}(\mathbf{x} + \mathbf{h})$  est proche de  $\mathbf{F}(\mathbf{x}) + \mathbf{L}(\mathbf{h})$  pour des  $\mathbf{h}$  petits (c'ad.  $\|\mathbf{h}\|$  est petit). Plus rigoureusement, soit  $\mathbf{R}_{\mathbf{F}}$  le terme d'erreur de l'approximation linéaire en  $\mathbf{x}$  :  $\mathbf{R}_{\mathbf{F}}(\mathbf{h}) = \mathbf{F}(\mathbf{x} + \mathbf{h}) - \mathbf{F}(\mathbf{x}) - \mathbf{L}(\mathbf{h})$ . À partir de quel moment l'approximation devient-elle "satisfaisante" ? Si  $\|\mathbf{R}_{\mathbf{F}}(\mathbf{h})\|$  décroît plus rapidement en norme vers 0 que  $\|\mathbf{h}\|$ , i.e. si

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\|\mathbf{R}_{\mathbf{F}}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

La raison intuitive qui se cache derrière ce critère est que si  $A$  est une matrice, alors  $\frac{\|A\mathbf{h}\|}{\|\mathbf{h}\|} \leq \sum_{i \in \overline{m}, j \in \overline{n}} a_{ij}^2$  (cf. exercice 26). Autrement dit, avec une décroissance aussi rapide que  $\|\mathbf{h}\|$ , n'importe quelle application linéaire ferait l'affaire, et, en particulier, n'importe quelle matrice  $B$  approchera la matrice  $A$ . D'où l'idée de demander un peu plus, comme dans l'équation ci-dessus.

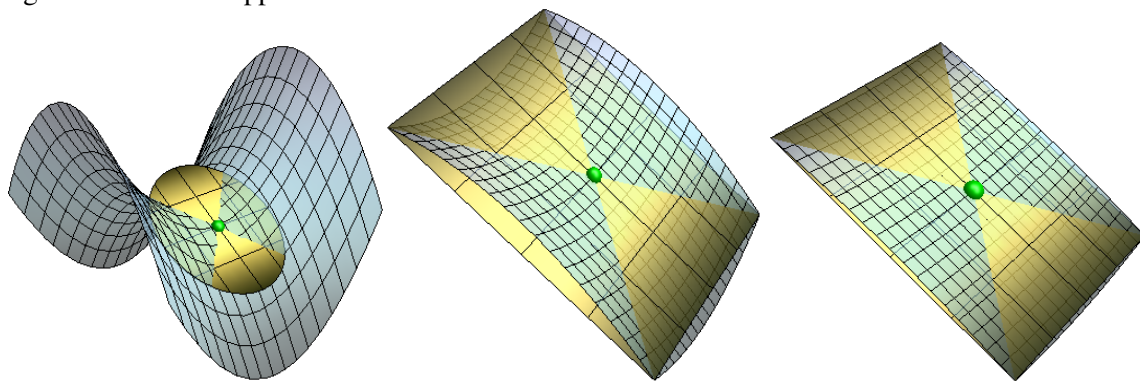
**Définition 3.2.4.** Soit  $S \subset \mathbb{R}^m$  et  $\mathbf{x}$  un point limite de  $S$ . Une application  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  est dite dérivable au point  $\mathbf{x}$  s'il existe une application linéaire  $\mathbf{L}$  telle que

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{F}(\mathbf{x} + \mathbf{h}) - \mathbf{F}(\mathbf{x}) - \mathbf{L}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}.$$

Une telle application linéaire  $\mathbf{L}$  est alors appelée dérivée [tout court] ou dérivée de Fréchet de  $\mathbf{F}$  en  $\mathbf{x}$ . ★

### 3.2 - Application linéaire et dérivabilité

Dans le graphique à gauche ci-dessous sont représentés, le graphe de la fonction  $F(x, y) = x^2 - y^2$  ainsi que l'image d'un disque par la dérivée (en tant qu'application linéaire) en un point  $\begin{pmatrix} a \\ b \end{pmatrix}$  (puis translatée par  $F(a, b)$  pour que ce point soit dans l'image). Les deux graphiques à droite sont des agrandissements successifs du premier ; plus le champ de vision se concentre près du point plus l'image de  $F$  et de son application linéarisée se confondent.



Si cette notion de dérivabilité porte parfois le nom de dérivabilité de Fréchet, c'est qu'il existe une autre notion, la dérivabilité de Gâteaux, qui est esquissée pour la curiosité du chaland en section 3.A. La notion de dérivée de Fréchet s'adapte bien au cadre infini-dimensionnel (où il faut remplacer  $\mathbf{L}$  par une application linéaire bornée) tout comme à l'analyse complexe (où  $\mathbf{L}$  devient alors une application linéaire sur  $\mathbb{C}$ , c'ad. une matrice à coefficients complexes).

Le fin grammairien qu'est le lecteur aura noté que la définition se formule avec "une telle application linéaire" plutôt que "l'application linéaire". Rien n'est gratuit :

#### **Théorème 3.2.5**

Soit  $U \subset \mathbb{R}^m$  un ouvert et  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application. Si  $\mathbf{F}$  est dérivable en  $\mathbf{x}$  alors sa dérivée est unique.

DÉMONSTRATION: Supposons qu'il existe deux matrices  $A_1, A_2$  et deux applications  $\mathbf{R}_1, \mathbf{R}_2 : U \rightarrow \mathbb{R}^n$  telles que

$$\begin{aligned} \mathbf{F}(\mathbf{x} + \mathbf{h}) - \mathbf{F}(\mathbf{x}) &= A_1 \mathbf{h} + \mathbf{R}_1(\mathbf{h}) \\ \mathbf{F}(\mathbf{x} + \mathbf{h}) - \mathbf{F}(\mathbf{x}) &= A_2 \mathbf{h} + \mathbf{R}_2(\mathbf{h}) \\ \text{et } \lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{R}_1(\mathbf{h})}{\|\mathbf{h}\|} &= \mathbf{0} = \lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{R}_2(\mathbf{h})}{\|\mathbf{h}\|} \end{aligned}$$

En prenant la différence des deux premières équations et en divisant par  $\|\mathbf{h}\|$  il apparaît que

$$(A_1 - A_2) \frac{\mathbf{h}}{\|\mathbf{h}\|} = \frac{\mathbf{R}_2(\mathbf{h}) - \mathbf{R}_1(\mathbf{h})}{\|\mathbf{h}\|}.$$

Puis, en posant  $\mathbf{y} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ ,  $\{t_k\}_{k \in \mathbb{Z}_{\geq 0}}$  une suite de réels convergeant vers 0 et  $\mathbf{h} = t_k \mathbf{y}$ , ceci devient

$$(A_1 - A_2) \frac{\mathbf{y}}{\|\mathbf{y}\|} = \frac{\mathbf{R}_2(t_k \mathbf{y}) - \mathbf{R}_1(t_k \mathbf{y})}{\|t_k \mathbf{y}\|}.$$

Par passage à la limite lorsque  $k$  tend vers l'infini des deux côtés de cette égalité,

$$\forall \mathbf{y} \in \mathbb{R}^m \setminus \{\mathbf{0}\}, (A_1 - A_2) \frac{\mathbf{y}}{\|\mathbf{y}\|} = \mathbf{0} \quad \Rightarrow \quad \forall \mathbf{y} \in \mathbb{R}^m, (A_1 - A_2) \mathbf{y} = \mathbf{0}$$

Ainsi, la matrice  $A_1 - A_2$  envoie tout  $\mathbb{R}^m$  vers 0, autrement dit  $A_1 = A_2$ . ■

Maintenant établie l'unicité de l'application linéaire qu'est la dérivée, il est possible de lui donner une notation. Pour  $\mathbf{F}$  une application linéaire dérivable en  $\mathbf{x}$ , la dérivée de  $\mathbf{F}$  en  $\mathbf{x}$  (en tant qu'application linéaire) sera notée  $d\mathbf{F}_{\mathbf{x}}$ .

EXERCICE 27: Soit  $U \subset \mathbb{R}^m$  un ouvert, et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application. Montrer que si  $\mathbf{F}$  est dérivable en  $\mathbf{x}$  alors  $\mathbf{F}$  est continue en  $\mathbf{x}$ .

### 3.3 Gradient et matrice de Jacobi

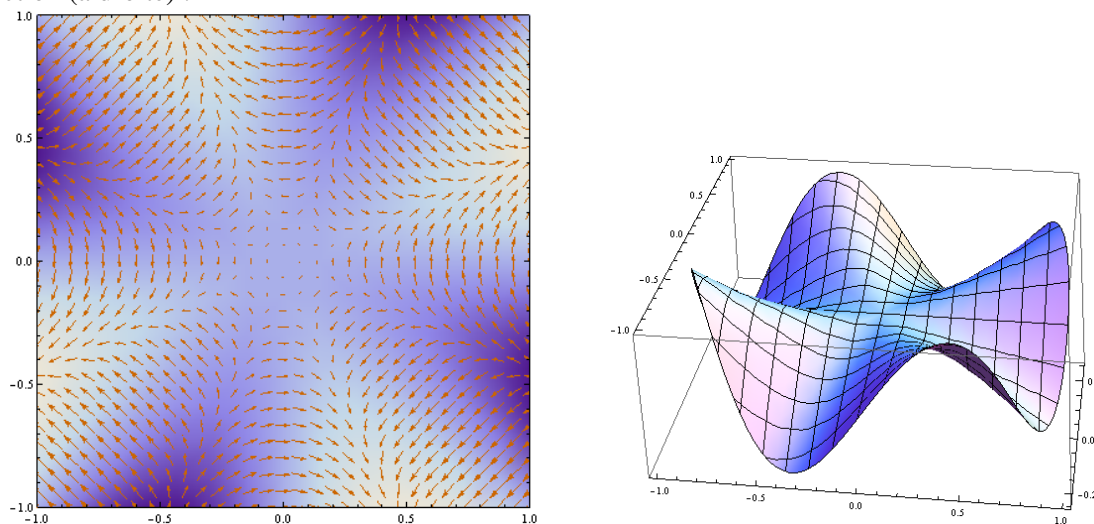
“Et quel est le lien entre cette dérivée un peu absconse et nos gentilles dérivées partielles ?” vous demandez-vous peut-être benoîtement. Il y en a un, qui porte deux noms, le gradient pour les fonctions, et la matrice de Jacobi (ou matrice jacobienne ou jacobienne [tout court], mais surtout il ne faut pas confondre avec le jacobien) pour les applications. Il s'agit foncièrement du même objet, mais le gradient possède plus de ramifications, et c'est ce qui motive cette séparation.

**Définition 3.3.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction telle que toutes les dérivées partielles existent en tout point de  $U$ . Le gradient de  $F$ , noté  $\nabla F$  ou  $(\text{grad} F)$ , est l'application (pour être plus pointu, le champ de vecteurs)  $\nabla F : \mathbb{R}^m \rightarrow \mathbb{R}^m$  défini par

$$\nabla F(\mathbf{x}) = \begin{pmatrix} D_1 F(\mathbf{x}) \\ D_2 F(\mathbf{x}) \\ \vdots \\ D_m F(\mathbf{x}) \end{pmatrix}. \quad \star$$

La définition ci-dessus insiste sur l'aspect champ de vecteurs du gradient. Lorsque toutes les dérivées partielles de  $F$  sont définies seulement en un point  $\mathbf{x}$ , il est aussi commode de dire que le gradient de  $F$  en  $\mathbf{x}$  est le point de  $\mathbb{R}^m$  donné par l'équation ci-dessus.

Voici une représentation du gradient de la fonction  $F(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$  (à gauche, avec un fond clair pour les grandes valeurs et un fond foncé pour les petites valeurs), ainsi que le graphe de la fonction (à droite) :



**Définition 3.3.2.** Soit  $U \subset \mathbb{R}^m$  un ouvert et  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application telle que toutes les dérivées partielles existent en  $\mathbf{y} \in U$ . La matrice de Jacobi ou matrice jacobienne de  $\mathbf{F}$  en  $\mathbf{y}$ , notée  $\mathbf{F}_*(\mathbf{y})$ , est la matrice de taille  $n \times m$  définie comme suit : pour  $i \in \bar{n}$  et  $j \in \bar{m}$ ,  $(\mathbf{F}_*(\mathbf{y}))_{ij} = D_j F_i(\mathbf{y})$ , ou

$$\mathbf{F}_*(\mathbf{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} F_1(\mathbf{x}) & \frac{\partial}{\partial x_2} F_1(\mathbf{x}) & \dots & \frac{\partial}{\partial x_m} F_1(\mathbf{x}) \\ \frac{\partial}{\partial x_1} F_2(\mathbf{x}) & \frac{\partial}{\partial x_2} F_2(\mathbf{x}) & \dots & \frac{\partial}{\partial x_m} F_2(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} F_n(\mathbf{x}) & \frac{\partial}{\partial x_2} F_n(\mathbf{x}) & \dots & \frac{\partial}{\partial x_m} F_n(\mathbf{x}) \end{pmatrix} \quad \star$$

### Théorème 3.3.3

Soit  $U \subset \mathbb{R}^m$  un ouvert et  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application. Si  $\mathbf{F}$  est dérivable au point  $\mathbf{x} \in U$  alors toutes les dérivées partielles de  $\mathbf{F}$  existent et sa dérivée est représentée matriciellement par la transposée de la matrice de Jacobi en  $\mathbf{x}$  i.e.

$$d\mathbf{F}_{\mathbf{x}}(\mathbf{v}) = \mathbf{F}_*(\mathbf{x}) \mathbf{v}$$

DÉMONSTRATION: Soit  $d\mathbf{F}_{\mathbf{x}}$  la dérivée de  $\mathbf{F}$  en  $\mathbf{x}$  et soit  $i \in \bar{n}$ . En utilisant la définition 3.2.4 de la dérivabilité, la fonction  $\mathbf{F}$  près de  $\mathbf{x}$  s'exprime par  $\mathbf{F}(\mathbf{x} + \mathbf{h}) - \mathbf{F}(\mathbf{x}) = d\mathbf{F}_{\mathbf{x}}(\mathbf{h}) + \mathbf{R}_{\mathbf{F}}(\mathbf{h})$ . Ainsi poser  $\mathbf{h} = h\mathbf{e}_i$  puis prendre la limite permet d'obtenir la définition 3.1.1 de la dérivée partielle donne

$$D_i \mathbf{F}(\mathbf{x}) = \lim_{h \rightarrow 0} \frac{\mathbf{F}(\mathbf{x} + h\mathbf{e}_i) - \mathbf{F}(\mathbf{x})}{h} = \lim_{h \rightarrow 0} \left( d\mathbf{F}_{\mathbf{x}}(\mathbf{e}_i) + \frac{\mathbf{R}_{\mathbf{F}}(h\mathbf{e}_i)}{h} \right) = d\mathbf{F}_{\mathbf{x}}(\mathbf{e}_i),$$

car  $\lim_{h \rightarrow 0} \frac{\mathbf{R}_{\mathbf{F}}(h\mathbf{e}_i)}{\|h\mathbf{e}_i\|} = 0$  et  $\|h\mathbf{e}_i\| = |h|$ . Si  $A$  est la matrice qui représente  $d\mathbf{F}_{\mathbf{x}}$ ,  $D_i \mathbf{F}(\mathbf{x}) = d\mathbf{F}_{\mathbf{x}}(\mathbf{e}_i) = A\mathbf{e}_i$  est la  $i^{\text{ème}}$  ligne de cette matrice. Or, il s'agit aussi de la  $i^{\text{ème}}$  colonne de la matrice de Jacobi. ■

**Exemple 3.3.4.** Soit  $U$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction dérivable sur  $U$ , alors la dérivée de  $\mathbf{F}$  en  $\mathbf{x}$  est la transposée de  $\nabla F(\mathbf{x})$ . D'où, pour  $\mathbf{h} \in \mathbb{R}^m$  petit, la variation de  $\mathbf{F}$  entre  $\mathbf{x}$  et  $\mathbf{x} + \mathbf{h}$  est d'environ  $\mathbf{h}(\nabla F(\mathbf{x}))^T = \nabla F(\mathbf{x}) \cdot \mathbf{h}$ .

Par exemple, la période d'un pendule est de  $P(L, g) = \sqrt{\frac{L}{g}}$  où  $L$  est la longueur du fil et  $g$  la gravité.  $\nabla P(L, g) = \left( \frac{1}{2L} \sqrt{\frac{L}{g}}, -\frac{1}{2g} \sqrt{\frac{L}{g}} \right)$ . En supposant une décroissance de  $g$  de 1% et une décroissance de  $L$  de 2%, la différence de période est d'environ  $\left( \frac{-2}{100}L, \frac{-1}{100}g \right) \cdot \left( \frac{1}{2L} \sqrt{\frac{L}{g}}, -\frac{1}{2g} \sqrt{\frac{L}{g}} \right) = \sqrt{\frac{L}{g}} \frac{1-2}{200}$ , c'est à dire une baisse de 0,5%. ♣

**Exemple 3.3.5.** L'application  $\mathbf{F} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  définie par

$$\mathbf{F}(R, r, \theta, \phi) = \begin{pmatrix} (R + r \cos \phi) \cos \theta \\ (R + r \cos \phi) \sin \theta \\ r \sin \phi \end{pmatrix}$$

a pour matrice de Jacobi en  $(R, r, \theta, \phi)$

$$\mathbf{F}_*(R, r, \theta, \phi) = \begin{pmatrix} \cos \theta & \cos \phi \cos \theta & -(R + r \cos \phi) \sin \theta & -r \sin \phi \cos \theta \\ \sin \theta & \cos \phi \sin \theta & (R + r \cos \phi) \cos \theta & -r \sin \phi \sin \theta \\ 0 & \sin \phi & 0 & r \cos \phi \end{pmatrix}.$$




**Théorème 3.3.6**

Soit  $U \subset \mathbb{R}^n$  un ouvert et  $F : U \rightarrow \mathbb{R}^n$  telle que  $\forall i \in \overline{m}, \forall \mathbf{x} \in U, D_i F(\mathbf{x})$  existent et, de plus, que les applications  $D_i F$  sont continues sur  $U$ . Alors  $F$  est dérivable en tout point de  $U$ .

DÉMONSTRATION: Afin de rendre la preuve plus lisible, elle ne sera ici présentée que dans le cas  $m = 2$  et  $n = 1$ . Ainsi, soit  $F : U \rightarrow \mathbb{R}$  la fonction considérée, soit  $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in U$  et soit  $\mathbf{h} = \begin{pmatrix} h \\ k \end{pmatrix} \in \mathbb{R}^2$  de petite norme (c'ad. de sorte que  $\mathbf{x} + \mathbf{h} \in U$ ). Le but de la preuve est de montrer que  $F(x+h, y+k) - F(x, y)$  s'approche d'un terme linéaire  $\mathbf{x} \cdot \mathbf{a}$  où  $\mathbf{a} = (D_1 F(\mathbf{x}), D_2 F(\mathbf{x}))$  et de termes d'erreurs  $R_F(h, k)$  qui tendent rapidement vers 0. Or

$$F(x+h, y+k) - F(x, y) = F(x+h, y+k) - F(x+h, y) + F(x+h, y) - F(x, y).$$

Par le théorème 3.1.6 de la valeur intermédiaire (pour les fonctions d'une variable),  $\exists s, t \in [0, 1]$  tels que

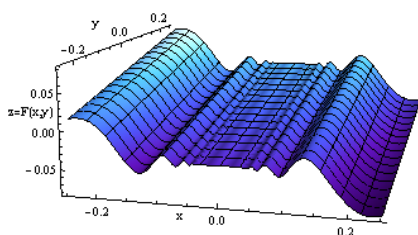
$$\begin{aligned} F(x+h, y+k) - F(x+h, y) &= k D_2 F(x+h, y+tk) \\ F(x+h, y) - F(x, y) &= h D_1 F(x+sh, y). \end{aligned}$$

Ainsi,  $R_F(h, k) = h(D_1 F(x+sh, y) - D_1 F(x, y)) + k(D_2 F(x+h, y+tk) - D_2 F(x, y))$ . Il ne reste plus qu'à vérifier que la décroissance est suffisante :

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{R_F(x, y)}{\|\mathbf{h}\|} = \lim_{\mathbf{h} \rightarrow \mathbf{0}} \left( \frac{h}{\|\mathbf{h}\|} (D_1 F(x+sh, y) - D_1 F(x, y)) + \frac{k}{\|\mathbf{h}\|} (D_2 F(x+h, y+tk) - D_2 F(x, y)) \right)$$

Or les deux fractions sont bornées ( $\frac{|h|}{\|\mathbf{h}\|} \leq 1$  et  $\frac{|k|}{\|\mathbf{h}\|} \leq 1$ ) et, par continuité, les différences de dérivées partielles ( $D_1 F(x+sh, y) - D_1 F(x, y)$  et  $D_2 F(x+h, y+tk) - D_2 F(x, y)$ ) tendent vers 0. La proposition 2.3.2.2 permet de conclure. ■

Les deux théorèmes précédents sont optimaux, c'ad. il n'est pas possible d'affaiblir les hypothèses, comme le montrent les exemples suivants. Si  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  est dérivable en  $\mathbf{x}$ , ses dérivées ne sont pas forcément continues :



**Exemple 3.3.7.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$F(x, y) = \begin{cases} x^2 \sin(\frac{1}{x}) + y & \text{si } x \neq 0 \\ y & \text{si } x = 0. \end{cases}$$



Alors  $D_1 F(x, y) = (2x - 1) \sin(\frac{1}{x})$  lorsque  $x \neq 0$  et  $D_1 F(0, y) = 0$  n'est pas continue en  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Pourtant  $F$  est dérivable en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ , sa dérivée est donnée par la matrice de taille  $1 \times 2$  suivante :  $(0, 1)$ . Les détails sont laissés au lecteur assidu (exercice suivant).

EXERCICE 28: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = \begin{cases} y + x^2 \sin \frac{1}{x} & \text{si } x \neq 0, \\ y & \text{si } x = 0. \end{cases}$$

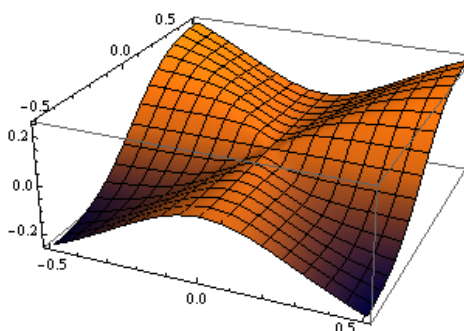
- Montrer que  $F$  est continue en  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ .
- Écrire les dérivées partielles de  $F$  en  $\mathbf{0}$ .
- Montrer que  $F$  est dérivable en  $\mathbf{0}$ .
- Montrer que  $D_1 F$  n'est pas continue en  $\mathbf{0}$ .

d'un autre côté, si les dérivées partielles ne sont pas continues,  $F$  n'est pas forcément dérivable :

**Exemple 3.3.8.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$F(x, y) = \begin{cases} \frac{xy^2}{x^2+y^2} & \text{si } (x, y) \neq \mathbf{0} \\ 0 & \text{si } x = y = 0. \end{cases}$$

Les dérivées partielles existent en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  (elles sont toutes deux nulles), mais  $F$  n'est pas dérivable en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . [Elle est par contre dérivable au sens de Gâteaux.]



La non-dérivabilité dans l'exemple ci-dessus provient de ce que le graphe contient 3 droites (distinctes) passant par 0 (comme visible sur le dessin). Or un plan (l'image d'une approximation linéaire) en 0 ne peut que contenir deux droites... voir exercice 35.

EXERCICE 29: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par

$$F(x, y) = y^2 + x^2 e^{\left(\frac{-1}{x^2+1}\right)}$$

Montrer que  $F$  est dérivable en  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . [Indice : Il y a plusieurs façons de faire. La méthode directe depuis la définition serait d'utiliser que  $e^t < 1$  lorsque  $t < 0$ .]

EXERCICE 30: Soit  $F(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$  l'application définie par

$$F(x, y) = \begin{cases} |xy|^{4/3} \sin\left(\frac{1}{xy}\right) & \text{si } xy \neq 0; \\ (0, 0) & \text{si } xy = 0. \end{cases}$$

- Montrer qu'elle est continue sur  $\mathbb{R}^2$ . [Vous pouvez utiliser, sans démonstration, que la fonction  $h : \mathbb{R} \rightarrow \mathbb{R}$  définie par  $h(t) = t \sin \frac{1}{t}$  lorsque  $t \neq 0$  et  $h(0) = 0$  est continue.]
- Montrer que les dérivées partielles existent en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  et donner leur valeur.
- Montrer qu'elle est dérivable en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . [Indice : trouver l'approximation linéaire candidate via la question précédente, puis utiliser que  $|x| \leq |x^2 + y^2|^{1/2}$  et  $|y| \leq |x^2 + y^2|^{1/2}$ .]

### 3.4 Dérivation composée

Les dérivées partielles et dérivées directionnelles ne sont qu'une infime partie des possibilités à disposition pour se ramener à une fonction d'une variable. En fait, soit  $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  une application et soit  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^m$  une courbe paramétrée, alors  $F \circ \gamma$  est une fonction d'une seule variable. Il est ainsi possible de s'amuser à définir non seulement des dérivées partielles (le long des axes de coordonnées) ou directionnelles (le long des droites) mais le long de presque n'importe quelle trajectoire.

Le but de cette section est plutôt d'essayer de calculer de telles dérivées à partir des outils déjà développés. Il ne s'agit en fait que de dérivation composée (ou dérivation en chaîne ou règle de Leibniz).

#### Théorème 3.4.1

Soit  $m, n, \ell \in \mathbb{Z}_{\geq 1}$ , soient  $U \subset \mathbb{R}^m$  et  $U' \subset \mathbb{R}^n$  des ouverts, soit  $\mathbf{F} : U \rightarrow U'$  une application dérivable en  $\mathbf{p} \in U$  et soit  $\mathbf{G} : U' \rightarrow \mathbb{R}^\ell$  une application dérivable en  $\mathbf{F}(\mathbf{p}) \in U'$ . Soit  $\mathbf{H} = \mathbf{G} \circ \mathbf{F}$  alors  $\mathbf{H}$  est dérivable en  $\mathbf{p}$  et les matrices de Jacobi sont reliées par le produit matriciel :

$$\mathbf{H}_*(\mathbf{p}) = \mathbf{G}_*(\mathbf{F}(\mathbf{p}))\mathbf{F}_*(\mathbf{p}).$$

DÉMONSTRATION: Soit  $M = \mathbf{G}_*(\mathbf{p}')$  la matrice représentant la dérivée de  $\mathbf{G}$  en  $\mathbf{p}' := \mathbf{F}(\mathbf{p})$  et  $L = \mathbf{F}_*(\mathbf{p})$  la matrice représentant la dérivée de  $\mathbf{F}$  en  $\mathbf{p}$ . Soit de plus  $\mathbf{R}_F(\mathbf{h}) = \mathbf{F}(\mathbf{p} + \mathbf{h}) - \mathbf{p}' - L\mathbf{h}$  et  $\mathbf{R}_G(\mathbf{v}) = \mathbf{G}(\mathbf{p}' + \mathbf{v}) - \mathbf{G}(\mathbf{p}') - M\mathbf{v}$ . La dérivabilité implique que  $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \mathbf{R}_F(\mathbf{h})/\|\mathbf{h}\| = \mathbf{0} = \lim_{\mathbf{v} \rightarrow \mathbf{0}} \mathbf{R}_G(\mathbf{v})/\|\mathbf{v}\|$ .

$$\begin{aligned} \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{H}(\mathbf{p} + \mathbf{h}) - \mathbf{H}(\mathbf{p}) - M L \mathbf{h}}{\|\mathbf{h}\|} &= \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{G}(\mathbf{F}(\mathbf{p} + \mathbf{h})) - \mathbf{G}(\mathbf{p}') - M L \mathbf{h}}{\|\mathbf{h}\|} \\ &= \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{G}(\mathbf{p}' + L\mathbf{h} + \mathbf{R}_F(\mathbf{h})) - \mathbf{G}(\mathbf{p}') - M L \mathbf{h}}{\|\mathbf{h}\|} \\ &= \lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{M L \mathbf{h} + M \mathbf{R}_F(\mathbf{h}) + \mathbf{R}_G(L\mathbf{h} + \mathbf{R}_F(\mathbf{h})) - M L \mathbf{h}}{\|\mathbf{h}\|} \\ &= \lim_{\mathbf{h} \rightarrow \mathbf{0}} \left( M \frac{\mathbf{R}_F(\mathbf{h})}{\|\mathbf{h}\|} + \frac{\mathbf{R}_G(L\mathbf{h} + \mathbf{R}_F(\mathbf{h}))}{\|L\mathbf{h} + \mathbf{R}_F(\mathbf{h})\|} \frac{\|L\mathbf{h} + \mathbf{R}_F(\mathbf{h})\|}{\|\mathbf{h}\|} \right) \end{aligned}$$

Pour  $\|\mathbf{h}\|$  assez petit, il existe un  $K \in \mathbb{R}_{>0}$  tel que

$$\frac{\|L\mathbf{h} + \mathbf{R}_F(\mathbf{h})\|}{\|\mathbf{h}\|} \leq \frac{\|L\mathbf{h}\|}{\|\mathbf{h}\|} + \frac{\|\mathbf{R}_F(\mathbf{h})\|}{\|\mathbf{h}\|} \leq \|L\|_2 + K$$

car  $\frac{\mathbf{R}_F(\mathbf{h})}{\|\mathbf{h}\|} \rightarrow \mathbf{0}$ . De plus,  $\lim_{\mathbf{h} \rightarrow \mathbf{0}} (L\mathbf{h} + \mathbf{R}_F(\mathbf{h})) = \mathbf{0}$ , et donc par les propriétés des limites (cf. proposition 2.3.2.1) et la continuité de la norme,  $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{R}_G(L\mathbf{h} + \mathbf{R}_F(\mathbf{h}))}{\|L\mathbf{h} + \mathbf{R}_F(\mathbf{h})\|} = \mathbf{0}$ . Les propriétés de la limite (cf. proposition 2.3.2.2) permettent de conclure que la limite ci-dessus est 0, en particulier que  $LM$  est la matrice de la dérivée de  $\mathbf{H}$  en  $\mathbf{p}$ . Une transposition permet d'obtenir l'assertion sur les matrices de Jacobi. ■

**Remarque 3.4.2.** Soit  $\mathbf{G} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  une application et soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction dérivable. Certains se rappellent de la règle de dérivation en chaîne en écrivant  $F \circ \mathbf{G}(x, y) = F(G_1(x, y), G_2(x, y))$ ,

### 3.4 - Dérivation composée

puis en posant deux variables accessoires  $u := G_1(x, y)$  et  $v := G_2(x, y)$ . Ceci donne

$$\begin{aligned}\frac{\partial}{\partial x} F(u, v) &= \frac{\partial F}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial F}{\partial v} \frac{\partial v}{\partial x} \\ &= (D_1 F)(u, v)(D_1 G_1)(x, y) + (D_2 F)(u, v)(D_1 G_2)(x, y) \\ \frac{\partial}{\partial y} F(u, v) &= \frac{\partial F}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial F}{\partial v} \frac{\partial v}{\partial y} \\ &= (D_1 F)(u, v)(D_2 G_1)(x, y) + (D_2 F)(u, v)(D_2 G_2)(x, y).\end{aligned}$$

En dimension supérieure la procédure est la même : si  $\mathbf{G} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  et  $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$ , et si pour  $i \in \overline{n}$ ,  $u_i = G_i(\mathbf{x})$ , alors, pour  $j \in \overline{m}$  et  $k \in \overline{\ell}$ ,

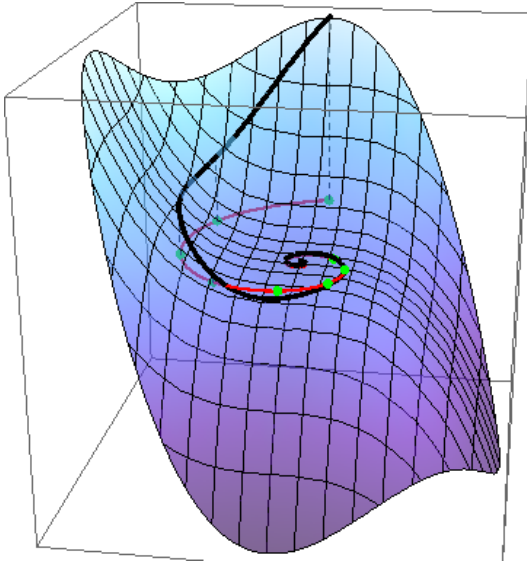
$$\frac{\partial}{\partial x_j} F_k(\mathbf{u}) = \sum_{i \in \overline{n}} \frac{\partial F_k}{\partial u_i} \frac{\partial u_i}{\partial x_j}.$$

Cette écriture est souvent préférée (car plus facile à retenir) ; l'essentiel étant de ne pas se tromper en oubliant la dépendance d'une variable accessoire sur une des variables d'origine. ♠

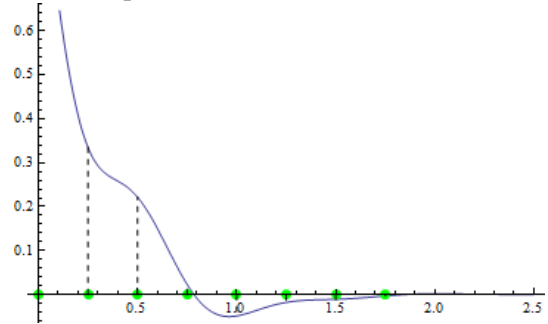
**Exemple 3.4.3.** Soit  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  la courbe paramétrée  $\gamma(t) = \begin{pmatrix} e^{-t} \cos t \\ e^{-t} \sin t \end{pmatrix}$  et soit  $F : \mathbb{R}^2 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$  définie par  $F(x, y) = x^3 + y^3$ . En calculant d'abord  $F_*(x, y) = (3x^2, 3y^2)$  et  $\gamma_*(t) = \begin{pmatrix} -e^{-t}(\cos t + \sin t) \\ e^{-t}(\cos t - \sin t) \end{pmatrix}$ , il apparaît que si  $h = F \circ \gamma$

$$h'(t) = F_*(\gamma(t))\gamma_*(t) = -3e^{-3t}(\cos^3 t + \cos^2 t \sin t - \cos t \sin^2 t + \sin^3 t)$$

Graphiquement, ceci peut se représenter comme suit : on trace d'abord le graphe de  $F$  (ici fait seulement pour  $(x, y) \in B_1(0, 0)$ ) puis dans le plan  $xy$  l'image de la courbe  $\gamma$  (ici fait pour  $t > 0$ , en rouge, avec des points en vert pour marquer des intervalles de longueur fixée).



Verticalement dans le graphe se trouve la valeur associée au point en bas, ceci trace une courbe dans le graphe (en noir). Le tout peut se remettre dans un graphe normal pour avoir la fonction composée.



♣

**Exemple 3.4.4.** Soit  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  définie par  $\mathbf{F}(r, \theta) = (\text{ch } r \cos \theta, \text{ch } r \sin \theta, \text{sh } r)$ , soit  $\mathbf{G}(x, y, z) = (x^2 + y^2 - z^2, x^2 + y^2 + z^2)$  et soit  $\mathbf{H} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  définie par  $\mathbf{H} = \mathbf{G} \circ \mathbf{F}$ . Composante par composante, la dérivée en chaîne peut s'évaluer en posant  $x = \text{ch } r \cos \theta$ ,  $y = \text{ch } r \sin \theta$  et  $z = \text{sh } r$ , puis par exemple

$$\frac{\partial H_1}{\partial r}(r, \theta) = \frac{\partial}{\partial r}(x^2 + y^2 - z^2) = 2x \frac{\partial}{\partial r} x + 2y \frac{\partial}{\partial r} y - 2z \frac{\partial}{\partial r} z = 2 \text{ch } r \text{sh } r - 2 \text{sh } r \text{ch } r = 0.$$

Sinon, les matrices de Jacobi sont

$$\mathbf{F}_*(r, \theta) = \begin{pmatrix} \operatorname{sh} r \cos \theta & -\operatorname{ch} r \sin \theta \\ \operatorname{sh} r \sin \theta & \operatorname{ch} r \cos \theta \\ \operatorname{ch} r & 0 \end{pmatrix} \text{ et } \mathbf{G}_*(x, y, z) = \begin{pmatrix} 2x & 2y & -2z \\ 2x & 2y & 2z \end{pmatrix}.$$

Elles se multiplient pour donner

$$\mathbf{H}_*(r, \theta) = \mathbf{G}_*(\mathbf{F}(r, \theta))\mathbf{F}_*(r, \theta) = \begin{pmatrix} 0 & 0 \\ 4\operatorname{ch} r \operatorname{sh} r & 0 \end{pmatrix} \quad \clubsuit$$

### 3.4.i Un théorème d'Euler sur les fonctions homogènes

Voici un corollaire célèbre des dérivées composées. Soit  $U \subset \mathbb{R}^m$  un ouvert, une fonction  $F : U \rightarrow \mathbb{R}$  est dite (positivement) homogène de degré  $k \in \mathbb{R}$ , si  $\forall a \in \mathbb{R}_{>0}, F(a\mathbf{x}) = a^k F(\mathbf{x})$ .

**Exemple 3.4.5.** Là où les formules suivantes ne comportent pas de division par 0, les fonctions  $F$  définies par  $F(x, y, z) =$

$$\frac{y}{\sqrt{x^2 + z^2}}, \quad z, \quad \frac{x + y + z}{x^2 + y^2 + z^2}, \quad \frac{y^2}{x}, \quad \log \left| \frac{3x - 5y}{2y - z} \right|, \quad \frac{3zx^2 + 4y^3}{x(y - 2z)} \quad \text{et} \quad \frac{xy(x^2 - y^2)}{x^2 + y^2}$$

sont homogènes de degré (respectivement) 0, 1, -1, 1, 0, 2 et 2. ♣

Si une fonction homogène est définie sur  $U$ , elle s'étend naturellement au cône sur  $U$  (l'ensemble  $\hat{U} := \cup_{t \in \mathbb{R}_{>0}} tU$ ) par homogénéité.

**EXERCICE 31:** Soit  $F$  une fonction homogène de degré  $k \geq 1$  définie sur un ouvert  $U \subset \mathbb{R}^m$  tel que  $U = \hat{U}$  et  $\mathbf{0} \in U$ . Si de plus  $F(\mathbf{0}) = 0$ , montrer alors que  $\forall \mathbf{u} \in U$  la dérivée directionnelle  $D_{\mathbf{u}}F(\mathbf{0}, 0)$  existe toujours et qu'elle est forcément nulle si  $k > 1$ .

**EXERCICE 32:** Soit  $F$  une fonction homogène de degré  $k = 1$  définie sur  $\mathbb{R}^2$ . Si de plus  $F(\mathbf{0}) = 0$ , montrer alors que s'il existe trois points  $(x_1, y_1), (x_2, y_2)$  et  $(x_3, y_3)$  sur le cercle (i.e.  $x_i^2 + y_i^2 = 1$ ) tels que  $i \neq j \Rightarrow |F(x_i, y_i)| \neq |F(x_j, y_j)|$  alors  $F$  n'est pas dérivable en  $(0, 0)$ .

#### Corollaire 3.4.6 (Théorème d'Euler pour les fonctions homogènes)

Soit  $U \subset \mathbb{R}^m$  un ouvert et  $F$  une fonction homogène de degré  $k$ , alors,  $\forall \mathbf{x} \in U$ ,

$$kF(\mathbf{x}) = \mathbf{x} \cdot \nabla F(\mathbf{x}) = \sum_{i=1}^m x_i D_i F(\mathbf{x}).$$

**DÉMONSTRATION:** Si  $\mathbf{x} \in U$  alors pour  $t \in \mathbb{R}_{>0}$  suffisamment proche de 1,  $t\mathbf{x} \in U$  (car  $d(\mathbf{x}, t\mathbf{x}) = \|\mathbf{x} - t\mathbf{x}\| = |1 - t| \|\mathbf{x}\|$ ). En prenant  $f(t) = F(t\mathbf{x})$  (définie et dérivable pour  $t$  proche de 1) s'obtient par dérivation en chaîne

$$f'(t) = \mathbf{x} \cdot \nabla F(t\mathbf{x}).$$

En effet, la courbe paramétrée  $\mathbf{G}(t) = t\mathbf{x}$  a pour matrice de Jacobi une matrice  $m \times 1$  (i.e. un vecteur) dont la  $i^{\text{ème}}$  entrée est  $x_i$  (la dérivation se fait par rapport à  $t$ ). De plus, la matrice de Jacobi de  $F(\mathbf{x})$  est la transposée de son gradient.

D'autre part, par homogénéité,  $f(t) = t^k F(\mathbf{x})$  et donc  $f'(t) = kt^{k-1} F(\mathbf{x})$ . Le résultat est obtenu en identifiant les deux expressions de  $f'(t)$  et en posant  $t = 1$ . ■

**Remarque 3.4.7.** Soit  $F : \mathbb{R}^m \setminus \{0\} \rightarrow \mathbb{R}$  une fonction homogène de degré 0. Supposons que  $\exists r \in \mathbb{R}_{>0}$  tel que  $F$  est continue sur  $\partial B_r(\mathbf{0})$ . Alors  $F$  est bornée. En effet,  $\partial B_r(\mathbf{0})$  est compact donc  $F$  y admet un maximum disons  $F(\mathbf{u})$ . Par homogénéité  $F(\mathbf{x}) = F(\frac{\mathbf{x}}{\|\mathbf{x}\|}r) \leq F(\mathbf{u})$ . ♠

EXERCICE 33: Soit  $F$  une fonction homogène de degré  $k > 0$  définie et continue sur  $\mathbb{R}^m$ . Montrer que  $\lim_{\mathbf{x} \rightarrow \mathbf{0}} F(\mathbf{x}) = 0$ . [Indice : utiliser le fait que  $F$  est essentiellement déterminée par sa valeur sur un compact.]

EXERCICE 34: Soit  $F$  une fonction homogène de degré  $k \in \mathbb{R}$  définie sur  $\{0\} \cup U$  où  $U \subset \mathbb{R}^m$  est un ouvert non-vide tel que  $U = \hat{U}$ .

Montrer que si  $k \leq 0$ , alors soit  $F$  est constante (et  $k = 0$ ) soit  $F$  n'est pas continue en  $\mathbf{0}$ .

EXERCICE 35: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction homogène de degré 1 (c'ad.  $\forall t \in \mathbb{R}_{>0}, F(t\mathbf{x}) = tF(\mathbf{x})$ ) telle que  $F(\mathbf{0}) = 0$ .

- Montrer que la dérivée directionnelle de  $F$  en  $\mathbf{0}$  existe pour toute direction.
- Le graphe de  $F$  est le sous-ensemble défini par  $\Gamma_F = \{(x, y, z) \in \mathbb{R}^3 \mid F(x, y) = z\}$ . Soit  $\mathbf{0} \neq \mathbf{p} \in \Gamma_F$ , montrer que  $\Gamma_F$  contient la demi-droite commençant en  $\mathbf{0} \in \mathbb{R}^3$  et passant par  $\mathbf{p}$ .
- Si  $F$  est dérivable en  $\mathbf{0}$ , montrer que le graphe de  $F$  est un plan.  
[Indice : la dérivée est une application linéaire qui approche "bien" la fonction, le graphe de la dérivée est un plan, or s'il existe une demi-droite qui n'est pas dans ce plan, l'approximation est "mauvaise".]
- Donner un exemple d'une fonction  $F$  non-dérivable en  $\mathbf{0}$  mais dont les dérivées directionnelles existent en  $\mathbf{0}$  peu importe la direction (et qui n'est pas une combinaison linéaire des fonctions déjà données précédemment). Expliquer (brièvement) pourquoi votre exemple satisfait les critères demandés.

## 3.5 Dérivées d'ordres supérieurs

A priori, la question des dérivées d'ordres supérieurs semble anodine. D'une fonction dérivable se déduisent ses dérivées partielles qui sont de nouveaux des fonctions, dérivables ou peut-être pas. Il n'y a en tout cas pas d'obstacle majeur à considérer des objets comme  $D_2 D_1 F$  ou  $D_1 D_1 F$  ou encore  $D_1 D_2 D_1 F$  dans la mesure où les limites sont définies. Les dérivées d'ordres supérieurs interviennent naturellement en physique (e.g. équation de la chaleur ou équation des ondes) et l'étude des fonctions de plusieurs variables (cf. en une variable) compte aussi parmi les motivations pour les considérer. Il n'est par contre pas aussi évident de savoir quand  $D_1 D_2 F = D_2 D_1 F$  ; en témoigne l'exemple suivant.

**Exemple 3.5.1.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction (homogène de degré 2) définie par

$$F(x, y) = \begin{cases} 0 & \text{si } x = y = 0 \\ \frac{xy(x^2 - y^2)}{x^2 + y^2} & \text{sinon.} \end{cases}$$

Un calcul direct révèle que

$$D_1 F(x, y) = \begin{cases} 0 & \text{si } x = y = 0 \\ \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & \text{sinon,} \end{cases} \quad \text{et } D_2 F(x, y) = \begin{cases} 0 & \text{si } x = y = 0 \\ \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & \text{sinon.} \end{cases}$$

L'évaluation de la limite donne ensuite  $D_2 D_1 F(0, 0) = -1$  et  $D_1 D_2 F(0, 0) = 1$ . ♣

Le résultat principal de cette section donne une condition sur  $F$  pour laquelle l'ordre des dérivées peut-être inversé.

**Définition 3.5.2.** Soit  $U \subset \mathbb{R}^m$  un ouvert et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application.  $\mathbf{F}$  est dite de classe

- $\mathcal{C}^0$  sur  $U$  si elle est continue sur  $U$ .
- $\mathcal{C}^1$  sur  $U$  si toutes ses dérivées sont continues sur  $U$ .
- $\mathcal{C}^k$ , où  $k \in \mathbb{Z}_{\geq 2}$ , si  $\forall \ell \in \bar{k}$  et pour tout  $i_1, i_2, \dots, i_\ell \in \bar{m}$ ,  $D_{i_1} D_{i_2} \cdots D_{i_\ell} \mathbf{F}$  est continue sur  $U$ .
- $\mathcal{C}^\infty$  si,  $\forall k \in \mathbb{Z}_{\geq 0}$  elle est de classe  $\mathcal{C}^k$

★

Le théorème suivant (donné avec des hypothèses très minimales) est le point de départ qui permettra d'inverser l'ordre des dérivations.

### Théorème 3.5.3

Soit  $U \subset \mathbb{R}^2$  un ouvert, soit  $F : U \rightarrow \mathbb{R}$  une fonction continue telle que  $D_1 F$ ,  $D_2 F$  et  $D_1 D_2 F$  existent en tout point de  $U$  que de plus  $D_1 D_2 F$  est continue en  $\begin{pmatrix} x \\ y \end{pmatrix} \in U$ . Alors  $D_2 D_1 F$  existe en  $\begin{pmatrix} x \\ y \end{pmatrix}$  et est égale à  $D_1 D_2 F(x, y)$ .

DÉMONSTRATION: Soit

$$G(h, k) = \frac{F(x+h, y+k) - F(x, y+k) - F(x+h, y) + F(x, y)}{hk}$$

définie pour  $h$  et  $k$  suffisamment petits (mais non-nuls). Soit  $a$  un réel tel que si  $V := ]-a, a[ \times ]-a, a[$  alors  $\forall (h, k) \in V, (x+h, y+k) \in U$  (ceci existe forcément car  $U$  est ouvert, voir remarque 2.2.6). Sur  $\pi_2(V) = ]-a, a[$ , soit définie, pour  $h \in \pi_1(V) = ]-a, a[$  fixé, la fonction  $f : \pi_2(V) \rightarrow \mathbb{R}$  par  $f(k) = F(x+h, y+k) - F(x, y+k)$ ; elle est dérivable en tout point. Le théorème de la valeur intermédiaire 3.1.6 implique qu'il existe  $t \in [0, 1]$  tel que  $f(k) - f(0) = kf'(tk)$ . Ceci permet d'écrire :  $G(h, k) = \frac{1}{hk}(f(k) - f(0)) = \frac{1}{h}f'(tk) = \frac{1}{h}(D_2 F(x+h, y+tk) - D_2 F(x, y+tk))$ . Une seconde application du théorème de la valeur intermédiaire 3.1.6 implique qu'il existe  $s \in [0, 1]$  tel que  $G(h, k) = D_1 D_2 F(x+sh, y+tk)$ . Or  $D_1 D_2 F$  est continue en  $\begin{pmatrix} x \\ y \end{pmatrix}$ , autrement dit  $\forall \varepsilon > 0, \exists \delta$ , tels que

$$\|(h, k)\| < \delta \Rightarrow \|G(h, k) - D_1 D_2 F(x, y)\| < \varepsilon$$

### 3.5 - Dérivées d'ordres supérieurs

En particulier, ceci reste vrai (pour le  $\varepsilon$  donné) si  $|h|$  et  $|k|$  sont  $< \delta/\sqrt{2}$ . Or  $\lim_{h \rightarrow 0} G(h, k) = \frac{1}{k} (D_1 F(x, y+k) - D_1 F(x, y))$ . Par conséquent,  $\forall \varepsilon, \exists \delta$ , tel que

$$|k| < \delta/\sqrt{2} \Rightarrow \left\| \frac{D_1 F(x, y+k) - D_1 F(x, y)}{k} - D_1 D_2 F(x, y) \right\| < \varepsilon.$$

Précisément, ceci veut dire que  $D_2 D_1 F$  existe en  $\binom{x}{y}$  et vaut  $D_1 D_2 F(x, y)$ . ■

Le reste n'est maintenant plus qu'une formalité.

#### Corollaire 3.5.4

Soit  $k \in \mathbb{Z}_{\geq 0}$ , soit  $F$  une application de classe  $\mathcal{C}^k$  sur  $U$ , soit  $\ell \in \bar{k}$ , soit  $i_1, i_2, \dots, i_\ell \in \bar{m}$  une suite finie de nombres entiers et soit  $i'_1, i'_2, \dots, i'_\ell \in \bar{m}$  une permutation de cette suite. Alors,  $\forall \mathbf{x} \in U$ ,

$$D_{i_1} D_{i_2} \cdots D_{i_\ell} F(\mathbf{x}) = D_{i'_1} D_{i'_2} \cdots D_{i'_\ell} F(\mathbf{x}).$$

DÉMONSTRATION: Tout d'abord si  $\ell = 1$  il n'y a rien à montrer ; il sera dorénavant supposé que  $\ell > 1$ . Les permutations sur  $\ell$  éléments sont engendrées par les permutations sur deux éléments consécutifs. Ainsi, il suffit de le montrer pour de telles permutations. De plus, il suffit de le faire sur  $i_1$  et  $i_2$  : en effet, si la permutation a lieu sur  $i_j$  et  $i_{j+1}$ , il est possible de se ramener au cas où  $\ell' = \ell - j + 1$ , i.e. de ne regarder que la suite  $D_{i_j} D_{i_{j+1}} \cdots D_{i_\ell}$ .

Or, comme  $\ell \leq k$ , la fonction  $G := D_{i_3} \cdots D_{i_\ell} F$  (lorsque  $\ell = 2$  il s'agit de  $F$  tout court) est de classe  $\mathcal{C}^{k-\ell+2}$  où  $k - \ell + 2 \geq 2$ . Si  $i_1 = i_2$ , il n'y a de nouveau rien à montrer. Si ces indices sont distincts, comme les autres coordonnées n'interviennent pas dans le calcul de  $D_{i_1}$  et  $D_{i_2}$  il est possible de considérer que  $G$  ne dépend que de deux variables et est de classe  $\mathcal{C}^2$ . Par le théorème 3.5.3, l'ordre des deux dérivées n'est pas important. ■

Lorsque que la fonction est de classe  $\mathcal{C}^k$  suffisamment grande,  $\ell \leq k$  dérivations partielles en la variable  $x_i$  seront notées  $D_i^\ell$  ou  $\frac{\partial^\ell}{\partial x_i^\ell}$ .

Soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^2$ .  $F$  est dite harmonique si  $\sum_{i \in \bar{m}} \frac{\partial^2}{\partial x_i^2} F(\mathbf{x}) = 0$ .

**Exemple 3.5.5.** Si  $F(x, y)$  est harmonique alors  $F(x^2 - y^2, 2xy)$  est harmonique. En effet, soit  $u = x^2 - y^2$  et  $v = 2xy$ ,

$$\begin{aligned} \frac{\partial}{\partial x} F(u, v) &= 2x(D_1 F)(u, v) + 2y(D_2 F)(u, v), \\ \frac{\partial}{\partial y} F(u, v) &= -2y(D_1 F)(u, v) + 2x(D_2 F)(u, v), \\ \frac{\partial^2}{\partial x^2} F(u, v) &= 2(D_1 F)(u, v) + 4x^2(D_1^2 F)(u, v) + 8xy(D_2 D_1 F)(u, v) + 4y^2(D_2^2 F)(u, v), \\ \frac{\partial^2}{\partial y^2} F(u, v) &= -2(D_1 F)(u, v) + 4y^2(D_1^2 F)(u, v) - 8xy(D_2 D_1 F)(u, v) + 4y^2(D_2^2 F)(u, v). \end{aligned}$$

Ainsi,  $\frac{\partial^2}{\partial x^2} F(u, v) + \frac{\partial^2}{\partial y^2} F(u, v) = 4(x^2 + y^2)((D_1^2 F)(u, v) + (D_2^2 F)(u, v)) = 0$ . ♣

EXERCICE 36: Montrer que si  $F(x, y)$  est harmonique alors  $F(x^3 - 3xy^2, 3x^2y - y^3)$  est harmonique.

EXERCICE 37: Montrer que si  $F(x, y)$  est harmonique alors  $F(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2})$  est harmonique.



EXERCICE 38: Montrer que la fonction  $G : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^2$  sur son domaine et définie par  $G(x, y) = \ln(x^2 + y^2)$  est harmonique. Montrer  $G(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}) = -G(x, y)$  et que  $G(x^3 - 3xy^2, 3x^2y - y^3) = 3G(x, y)$ .

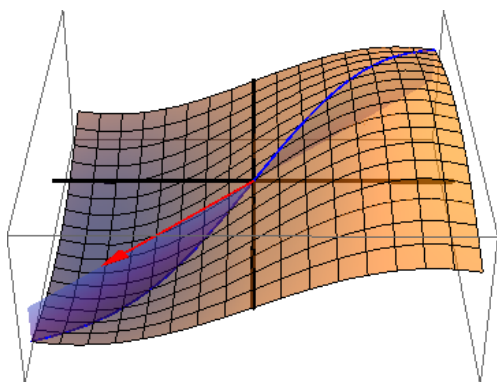
EXERCICE 39: Montrer que si  $F(x, y)$  est harmonique alors  $F(ax + by, bx - ay)$  est harmonique.

## 3.6 Interprétation géométrique quand $m = 1$ ou $n = 1$

L'objectif de cette section est de présenter quelques interprétations ou applications simples des dérivées, principalement du gradient.

### 3.6.i Dérivée directionnelle

Lorsque les dérivées partielles sont définies, il y a une grosse dose d'arbitraire de choisir les axes de coordonnées. En effet, plutôt que de poser  $f(t) = F(x_1 + t, x_2, \dots, x_m) = F(\mathbf{x} + t\mathbf{e}_1)$  pour obtenir une dérivée partielle, il est tout à fait possible de retomber sur le cas d'une variable, en regardant  $f(t) = F(\mathbf{x} + t\mathbf{y})$ . C'est l'idée de la dérivée directionnelle. Un changement de variable assez simple permet de voir que la norme de  $\mathbf{y}$ , ne joue pas réellement de rôle ; ainsi, un vecteur unitaire lui est préféré, c'est-à-dire  $\mathbf{u} \in \mathbb{R}^m$  tel que  $\|\mathbf{u}\| = 1$ .



Par exemple, dans le dessin ci-contre, est représenté le graphe d'une fonction  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Les deux axes de coordonnées sont tracés dans le plan  $xy$  en noir. Le vecteur en rouge pointe dans une autre direction et le long de cette droite apparaît en bleu l'aire sous le graphe d'une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Prendre la dérivée directionnelle (pour un vecteur unitaire) c'est regarder la dérivée de la fonction  $f$ .

**Définition 3.6.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert, et  $\mathbf{F} : U \rightarrow \mathbb{R}^n$ . La dérivée de  $\mathbf{F}$  en  $\mathbf{x} \in U$  dans la direction  $\mathbf{u} \in \mathbb{R}^m$  (où  $\|\mathbf{u}\| = 1$ ), notée  $D_{\mathbf{u}}\mathbf{F}(\mathbf{x})$ , est définie par

$$D_{\mathbf{u}}\mathbf{F}(\mathbf{x}) = \lim_{t \rightarrow 0^+} \frac{\mathbf{F}(\mathbf{x} + t\mathbf{u}) - \mathbf{F}(\mathbf{x})}{t} = \left. \frac{d}{dt} \mathbf{F}(\mathbf{x} + t\mathbf{u}) \right|_{t=0}$$

si la limite existe.



Le résultat suivant est laissé en exercice (cf. preuve du théorème 3.3.3).

EXERCICE 40: Soit  $U$  un ouvert de  $\mathbb{R}^m$ ,  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application dérivable en  $\mathbf{x}$ ,  $\mathbf{u} \in \mathbb{R}^m$  un vecteur (basé en  $\mathbf{x}$ ). Montrer que la dérivée directionnelle de  $\mathbf{F}$  dans la direction de  $\mathbf{u}$  est exprimée par  $D_{\mathbf{u}}\mathbf{F}(\mathbf{x}) = \mathbf{F}_*(\mathbf{x})\mathbf{u}$ . En particulier, montrer que, lorsque  $n = 1$ ,  $D_{\mathbf{u}}F(\mathbf{x}) = \mathbf{u} \cdot \nabla F(\mathbf{x})$ .

Ceci motive une notation accorte qui sera en particulier utile pour l'expression de la formule de Taylor-Lagrange (cf. section 5.1) :

$$(3.6.2) \quad D_{\mathbf{u}}F = (\mathbf{u} \cdot \nabla)F \quad \text{et} \quad \underbrace{D_{\mathbf{u}}D_{\mathbf{u}} \cdots D_{\mathbf{u}}F}_{k \text{ fois}} = (D_{\mathbf{u}})^k F = (\mathbf{u} \cdot \nabla)^k F.$$

L'existence de dérivées directionnelles en  $\mathbf{x}$  dans toutes les directions n'est pas une garantie suffisante pour la dérivabilité, cf. la fonction de l'exemple 2.4.4 ; le lecteur curieux est prié d'assouvir son appétence dans la section 3.A.

La dérivée directionnelle permet de donner une interprétation au gradient. En effet, si  $\nabla F(\mathbf{x}) \neq \mathbf{0}$ , alors  $\nabla F(\mathbf{x})$  vu comme vecteur basé en  $\mathbf{x}$  pointe dans la direction du plus grand accroissement en  $\mathbf{x}$  (voir les premiers graphiques de la section 3.3). Inversement,  $-\nabla F(\mathbf{x})$  pointe dans la direction où (depuis  $\mathbf{x}$ ) la fonction décroît le plus. Dans ces deux cas, l'information est uniquement la direction du vecteur ; sa longueur indique à quel point la croissance ou décroissance est rapide.

**Exemple 3.6.3.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = x^2 + y^2 = \|(x, y)\|^2$ , alors  $\nabla F(x, y) = (2x, 2y) = 2(x, y)$ . En chaque point  $(x, y) \in \mathbb{R}^2$ , la direction maximale de croissance est donnée par le vecteur  $2(x, y)$ , c'est-à-dire la direction qui s'éloigne de l'origine  $(0, 0)$ . Par contre,  $\nabla F(0, 0) = (0, 0)$ , donc (en ne considérant que le gradient) aucune conclusion ne peut-être tirée. ♣

Dans l'exemple précédent, le point où  $\nabla F$  s'annulait était aussi le point où  $F$  prenait sa plus petite valeur (voir la propriété de positivité de la norme No1 en section 1.3). Ce n'est pas un hasard, mais cette relation est plus subtile qu'il n'y paraît et sera approfondie au chapitre suivant.

EXERCICE 41: Soit  $F$  une fonction homogène de degré  $k \geq 1$  définie sur un ouvert  $U \subset \mathbb{R}^m$  tel que  $U = \hat{U}$  et  $\mathbf{0} \in U$ . Si  $F(\mathbf{0}) = 0$ , montrer alors que  $\forall \mathbf{u} \in U$  la dérivée directionnelle  $D_{\mathbf{u}}F(0, 0)$  existe toujours et qu'elle est forcément nulle si  $k > 1$ .

EXERCICE 42: Pour chacune des fonctions  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  de classe  $C^1$  suivantes, calculer son gradient, déterminer quand il est nul, et dessiner le champ de vecteurs donné par la direction du gradient  $\mathbf{G}(x, y) = \frac{\nabla F(x, y)}{\|\nabla F(x, y)\|}$  sur la région spécifiée.

- $F(x, y) = xy$  sur  $[-1, 1] \times [-1, 1]$  ;
- $F(x, y) = x^2(1 - x^2/2) + y^2(1 - y^2/2)$  sur  $[-2, 2] \times [-2, 2]$ . [Indice : regarder le signe du polynôme  $z - z^3$  pour  $z \in \mathbb{R}$ .]

EXERCICE 43: À l'exercice 23, il était démontré que la fonction  $T : (\mathbb{R} \setminus \{0\}) \times \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$T(x, y) = \begin{cases} \operatorname{arctg} \frac{y}{x} & \text{si } x > 0, \\ \pi + \operatorname{arctg} \frac{y}{x} & \text{si } x < 0, \end{cases}$$

possède des dérivées partielles définies partout sauf en  $(0, 0)$  (après extension [de ces dérivées] par continuité). Soit  $\mathbf{F}$  le champ de vecteurs donné par l'extension par continuité de  $\nabla T$  à  $\mathbb{R}^2 \setminus \{0\}$ .

- Montrer que  $\|\mathbf{F}(x, y)\| = (x^2 + y^2)^{-1}$ .
- Représenter graphiquement le champ de vecteurs  $\mathbf{F}(x, y)$  sur  $\{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 4\}$ .

### 3.6.ii Potentiel

En l'absence d'effets dispersifs (frottements, ...), il est en physique souvent très utile de regarder les forces comme découlant d'un potentiel. La force gravitationnelle que crée une masse ou la force électrique causée par une charge se représentent comme un champ de vecteurs  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ , c'est-à-dire qu'en chaque point de l'espace  $\mathbf{x} \in \mathbb{R}^3$ , il y a un vecteur  $\mathbf{v} \in \mathbb{R}^3$  (basé en  $\mathbf{x}$ ) qui décrit la direction et l'intensité de cette force au point  $\mathbf{x}$ .

Soit  $S \subset \mathbb{R}^3$  un ouvert. Un champ de vecteurs  $\mathbf{F} : S \rightarrow \mathbb{R}^3$  est dit conservatif s'il existe une fonction  $V : S \rightarrow \mathbb{R}$  dérivable telle que  $\mathbf{F}(x, y, z) = -\nabla V(x, y, z)$ .  $V$  est alors nommé un potentiel pour la force  $\mathbf{F}$  (l'ajout d'une constante transforme un potentiel en un nouveau potentiel).

**Exemple 3.6.4.** La force gravitationnelle que crée une masse ponctuelle en l'origine est le champ de vecteurs (défini sur  $\mathbb{R}^m \setminus \{\mathbf{0}\}$ ) donné par  $\mathbf{F}(\mathbf{x}) = -\frac{\mathbf{x}}{\|\mathbf{x}\|^3}$ . Il est possible de lui associer un potentiel, e.g.  $V(\mathbf{x}) = -\frac{1}{\|\mathbf{x}\|}$ . ♣

### 3.6.iii Vitesse

Une courbe paramétrée est une application  $\mathbb{R} \supset I \rightarrow \mathbb{R}^m$ . La manière la plus commode d'y penser est comme une application qui à un temps  $t \in I \subset \mathbb{R}$  donne la position d'une particule au temps  $t$ . Ainsi, une courbe paramétrée  $\gamma$  s'écrit  $\gamma(t) = (\gamma_1(t), \gamma_2(t), \dots, \gamma_m(t))^T$  où les  $\gamma_i$  sont les fonctions coordonnées.

Dans ce cas, il n'y a qu'une seule dérivée partielle de  $\gamma$ , et sa matrice de Jacobi n'est donc qu'une simple colonne. L'approximation linéaire que forme la dérivée n'est finalement qu'un vecteur ligne, rien d'autre que  $D_1\gamma$ . Par commodité, ce vecteur est plutôt noté  $\gamma'$ , i.e.

$$\gamma'(t) := D_1\gamma(t) = \gamma_*(t) = (\gamma'_1(t), \gamma'_2(t), \dots, \gamma'_m(t))^T.$$

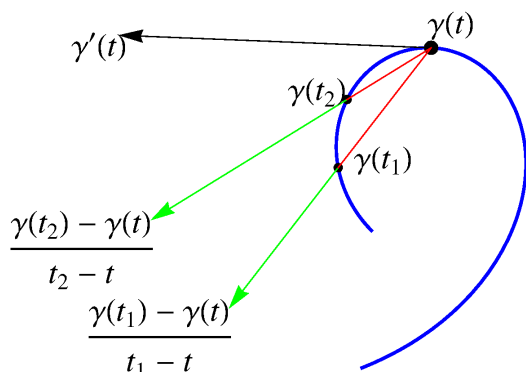
L'approximation linéaire est ainsi

$$\gamma(t + \varepsilon) \simeq \gamma(t) + \varepsilon\gamma'(t).$$

Comme prévu, c'est l'approximation par une droite (de vecteur directeur  $\gamma'(t)$ ). Une autre écriture de  $\gamma'$  est

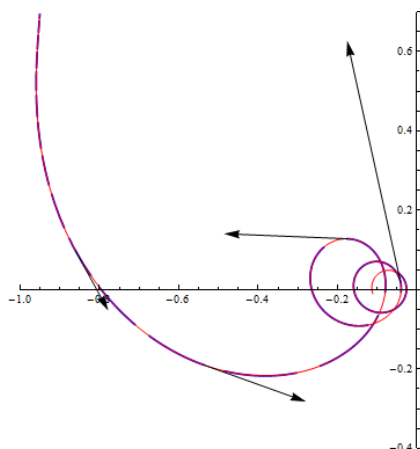
$$\gamma'(t) = \lim_{t' \rightarrow 0} \frac{\gamma(t + t') - \gamma(t)}{t'}.$$

Autrement dit, c'est la vitesse instantanée (plutôt le vecteur vitesse) de la particule au temps  $t$ .



Effectivement, le déplacement entre deux temps  $t$  et  $t + \varepsilon$  est donné par  $\gamma(t + \varepsilon) - \gamma(t)$  et l'intervalle de temps nécessaire à le parcourir  $\varepsilon$ . Ainsi la vitesse moyenne (et la direction) est naturellement définie par le rapport  $\frac{\gamma(t + \varepsilon) - \gamma(t)}{\varepsilon}$  (pour  $\varepsilon$  petit, voir ce que donne de l'approxima-

tion linéaire).



Dans le dessin ci-dessus, la courbe est tracée avec des changements de texture pour indiquer la vitesse, plus la partie où le trait est aminci est longue plus la vitesse est grande. Les vecteurs vitesse sont aussi représentés en quelques points choisis.

## 3.7 Espace(s) tangent(s)

Le but de cette section n'est plus d'approcher une fonction par une fonction linéaire (en fait, affine), mais d'approcher des ensembles par des ensembles affins. Le terme linéaire a un bon comportement sous/avec l'opération  $aX + bY$ , et le terme affine diffère de linéaire par une translation.

“Ensemble affine” (tout comme ensemble linéaire) peut ici être compris de plusieurs manières différentes, et cette multitude de perceptions se reflétera dans autant de manières de vouloir approcher un ensemble. Voici donc trois façons de définir un ensemble affine :

- EA1- Un ensemble affine est la solution d'une équation affine, *c'àd.* l'ensemble des  $\mathbf{x} \in \mathbb{R}^m$  tels que  $A(\mathbf{x} - \mathbf{b}) = \mathbf{0}$  où  $A$  est une matrice  $n \times m$  et  $\mathbf{b} \in \mathbb{R}^m$ . Ici,  $S = \{\mathbf{x} \in \mathbb{R}^m \mid A\mathbf{x} = \mathbf{b}'\}$  est de dimension  $\max(0, m - n)$  à  $m$  (où  $\mathbf{b}' = A\mathbf{b} \in \text{Im}A$ ).
- EA2- Un ensemble affine est l'ensemble des points  $\mathbf{x}$  obtenus d'un point  $\mathbf{a}$  et de  $\ell$  vecteurs  $\mathbf{v}_{(1)}, \mathbf{v}_{(2)}, \dots, \mathbf{v}_{(\ell)}$ . Ici,  $S = \{\mathbf{x} \in \mathbb{R}^m \mid \exists c_1, c_2, \dots, c_\ell \in \mathbb{R}, \mathbf{x} = \mathbf{a} + \sum_{i \in \bar{\ell}} c_i \mathbf{v}_{(i)}\}$  est de dimension 0 à  $\min(\ell, m)$  (selon le nombre de vecteurs linéairement indépendants).
- EA3- Un ensemble affine est l'image d'une application affine, *c'àd.* l'ensemble des  $\mathbf{x}$  tels qu'il existe  $\mathbf{y} \in \mathbb{R}^\ell$  pour lequel  $\mathbf{x} = B\mathbf{y} + \mathbf{a}$  (où  $B$  est une matrice  $m \times \ell$  et  $\mathbf{a} \in \mathbb{R}^m$  sont les deux ingrédients qui définissent l'ensemble affine). Ici,  $S := \text{Im}B + \mathbf{a} := \{\mathbf{x} \in \mathbb{R}^m \mid \exists \mathbf{y} \in \mathbb{R}^\ell \text{ avec } \mathbf{x} = \mathbf{a} + B\mathbf{y}\}$ . Sa dimension va de 0 à  $\min(\ell, m)$ .

Une des premières étapes d'un cours d'algèbre linéaire est de s'empresse de montrer qu'il n'y a pas de différences entre ces descriptions. Ceci est plutôt fait au niveau des applications linéaires ou des espaces vectoriels (*i.e.* les ensembles linéaires), et le passage à “affine” se fait en vérifiant que l'ajout d'une translation ne dérange pas grand chose (à ce niveau). La tâche sera plus ardue ici (de fait, il faudra attendre au chapitre prochain pour la terminer).

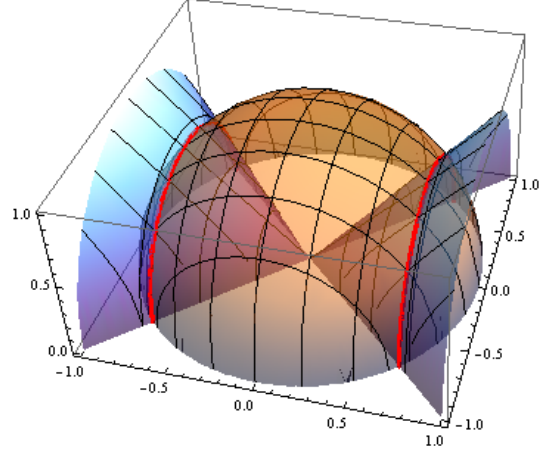
### 3.7.i Tangence aux ensembles de niveau

Tout d'abord, tentons de réinterpréter EA1 dans le contexte des applications (générales, *i.e.* pas forcément affines). Dans EA1, l'ensemble affine est l'ensemble des “zéros” d'une application affine.

Il est pour le moins logique de regarder l'ensemble des "zéros" d'une application. Autrement dit, de regarder un ensemble de niveau d'une application.

Ainsi, soit  $U \subset \mathbb{R}^m$  un ouvert, soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application, soit  $\mathbf{c} \in \mathbb{R}^n$  et soit  $S_{\mathbf{c}} = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{F}(\mathbf{x}) = \mathbf{c}\}$  l'ensemble de niveau de  $\mathbf{F}$  associé à  $\mathbf{c}$ . Comme il y a plusieurs fonctions coordonnées contenues dans l'application  $\mathbf{F}$ , remarquons que, si  $F_i$  est la  $i^{\text{ème}}$  fonction coordonnée et  $c_i$  la  $i^{\text{ème}}$  composante de  $\mathbf{c}$  (où  $i \in \overline{n}$ ), alors, si  $\mathbf{x} \in S_{\mathbf{c}}$ ,  $\mathbf{x} \in F_i^{-1}(c_i)$ . Ainsi, il est possible de ré-écrire

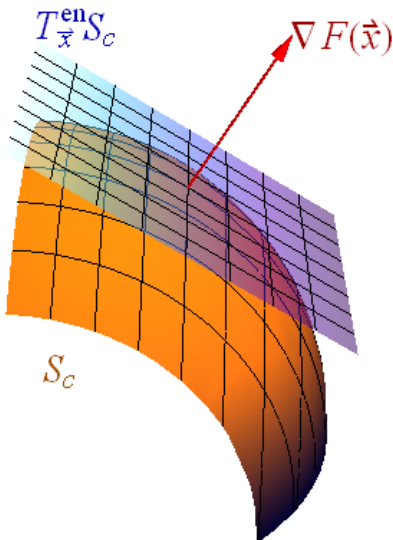
$$S_{\mathbf{c}} = \bigcap_{i \in \overline{n}} F_i^{-1}(c_i).$$



Par exemple, ci-dessus sont représentés deux ensembles de niveaux de  $\mathbb{R}^3$  (en fait pour la clarté, seuls les points où  $z > 0$  ont été pris en compte) : la sphère (ensemble du niveau 0 de  $F_1(x, y, z) = x^2 + y^2 + z^2 - 1$ ) et le cône double (ensemble du niveau 0 de  $F_2(x, y, z) = x^2 + y^2 - z^2$ ). Leur intersection (deux cercles) est la surface de niveau  $S_{(0,0)}$  de l'application  $\mathbf{F}(x, y, z) = \begin{pmatrix} F_1(x, y, z) \\ F_2(x, y, z) \end{pmatrix}$ . Ainsi, fidèles à notre habitude simplificatrice, commençons par le cas où  $\mathbf{F}$  n'est qu'une fonction  $F : \mathbb{R}^m \rightarrow \mathbb{R}$  (et  $\mathbf{c} = c \in \mathbb{R}$ ).

Dans ce cas, les directions  $\mathbf{u}$  qui sont orthogonales à  $\nabla F(\mathbf{x})$  en un point  $\mathbf{x} \in S_c$ , sont des directions où il n'y a (infinitésimalement) pas de variation. Autrement dit, il s'agit, dans la représentation de l'ensemble de niveau, de directions (depuis  $\mathbf{x}$ ) où la valeur de  $F$  change peu. C'est là la première idée intuitive pour l'ensemble affín qui approchera l'ensemble  $S_c$  : proche de  $\mathbf{x}$ ,  $F$  est approchée par une application affín  $A$  donnée par  $\mathbf{v} \mapsto A(\mathbf{v}) = F(\mathbf{x}) + \mathbf{v} \cdot \nabla F(\mathbf{x})$ , et l'ensemble affín sera l'ensemble de niveau de  $A$  :  $A^{-1}(c)$ . Une simple manipulation des égalités donne que  $A^{-1}(c) = \{\mathbf{v} \in \mathbb{R}^m \mid \mathbf{v} \cdot \nabla F(\mathbf{x}) = 0\}$ .

Pour passer à la dimension supérieure, le raisonnement est le même (l'argument de la dérivée directionnelle en moins) :  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  est approchée en  $\mathbf{x}$  par  $\mathbf{F}(\mathbf{x} + \mathbf{v}) \simeq \mathbf{F}(\mathbf{x}) + \mathbf{F}_*(\mathbf{x})\mathbf{v}$ . L'ensemble qui approchera  $S_{\mathbf{c}}$  en  $\mathbf{x}$  sera donné par les vecteurs tels que  $\mathbf{A}(\mathbf{v}) := \mathbf{F}(\mathbf{x}) + \mathbf{F}_*(\mathbf{x})\mathbf{v} = \mathbf{c}$  où, en fait,  $\mathbf{c} = \mathbf{F}(\mathbf{x})$ . Ceci motive la définition suivante :



**Définition 3.7.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert,  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application,  $\mathbf{c} \in \mathbb{R}^n$ ,  $S_{\mathbf{c}} = \mathbf{F}^{-1}(\mathbf{c})$  un ensemble de niveau (non-vidé) et  $\mathbf{x} \in S_{\mathbf{c}}$ . Alors l'espace tangent à  $S_{\mathbf{c}}$  est l'espace vectoriel défini

par

$$T_{\mathbf{x}}^{en} S_c = \{\mathbf{v} \in \mathbb{R}^m \mid \mathbf{F}_*(\mathbf{x})\mathbf{v} = \mathbf{0}\}.$$

D'autre part, le plan tangent est l'espace affine

obtenu par translation de  $T_{\mathbf{x}}^{en} S_c$  de manière à ce qu'il contienne  $\mathbf{x}$  :  $P_{\mathbf{x}}^{en} S_c = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{F}_*(\mathbf{x})(\mathbf{y} - \mathbf{x}) = \mathbf{0}\}.$  ★

On peut d'ailleurs se rapporter aux dessins qui se situent entre la définition 3.2.4 et le théorème 3.2.5 : de loin, le plan tangent n'a pas nécessairement l'air proche de la surface de niveau, il faut aller voir de très près pour comprendre qu'il y a tangence.

Il est important de remarquer qu'un ensemble fixé peut-être réalisé comme l'ensemble de niveau de plusieurs fonction différente. Pour prendre un exemple simple, si  $S = F^{-1}(0)$ , et  $G(\mathbf{x}) := F(\mathbf{x})^2$ , alors  $S = G^{-1}(0)$ . De plus,  $\nabla G(\mathbf{x}) = 2F(\mathbf{x})\nabla F(\mathbf{x})$ , ce qui implique que l'espace tangent en tant qu'espace de niveau pourra être différent (aux points où  $F(\mathbf{x}) = 0$  mais  $\nabla F(\mathbf{x}) \neq \mathbf{0}$ ).

Tout d'abord (pour se placer par rapport à la terminologie de l'introduction de cette section) le  $T$  est un espace vectoriel, tandis que le  $P$  est un espace affine.

Ensuite,  $T_{\mathbf{x}}^{en} S_c$  peut se décrire de manière plus géométrique. En effet,  $\mathbf{F}_*(\mathbf{x})\mathbf{v}$  est le vecteur  $(\mathbf{v} \cdot \nabla F_1(\mathbf{x}), \mathbf{v} \cdot \nabla F_2(\mathbf{x}), \dots, \mathbf{v} \cdot \nabla F_n(\mathbf{x}))^T$ . Autrement dit, l'espace tangent à  $S_c$  en  $\mathbf{x}$  est l'espace des vecteurs orthogonaux aux gradients des fonctions coordonnées (évaluées en  $\mathbf{x}$ ). Cependant, l'espace tangent à  $F_i^{-1}(c_i)$  en  $\mathbf{x}$  est donné par l'orthogonalité à  $\nabla F_i(\mathbf{x})$ . Ainsi, l'espace tangent à l'intersection des surface de niveaux  $\cap_{i \in \bar{n}} F_i^{-1}(c_i)$  est l'intersection des espaces tangents à chaque  $F_i^{-1}(c_i)$ . Autrement dit, notre approximation linéaire semble bien marcher sous intersection.

Par définition, l'espace tangent  $T_{\mathbf{x}}^{en} S_c$  est un espace vectoriel. Selon la taille de l'espace linéaire engendré par les  $\nabla F_i(\mathbf{x})$  sa dimension est incluse entre  $m - n$  (si les  $\nabla F_i(\mathbf{x})$  sont linéairement indépendants) et  $m$  (s'ils sont tous nuls).

**Exemple 3.7.2.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = x^2 - y^2$ . Son gradient est  $\nabla F(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$ . À  $c \in \mathbb{R}$  est associé l'ensemble de niveau  $S_c = F^{-1}(c)$ .

Au point  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  dans l'ensemble de niveau  $S_0$ , la droite tangente est faite de vecteurs  $\mathbf{v} = \begin{pmatrix} u \\ v \end{pmatrix}$  (basés en  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ) satisfaisant  $u - v = 0$ . En tant que plan tangent, il s'agit des points  $\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0 \Leftrightarrow x = y$ .

En fait, pour  $c = 0$  l'ensemble de niveau est lui-même fait des deux droites  $x = \pm y$ . D'ailleurs, en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  le gradient s'annule ; en conséquence l'espace tangent à cet endroit (tout  $\mathbb{R}^2$ ) ne correspond pas trop à une approximation linéaire de  $S_0$ .

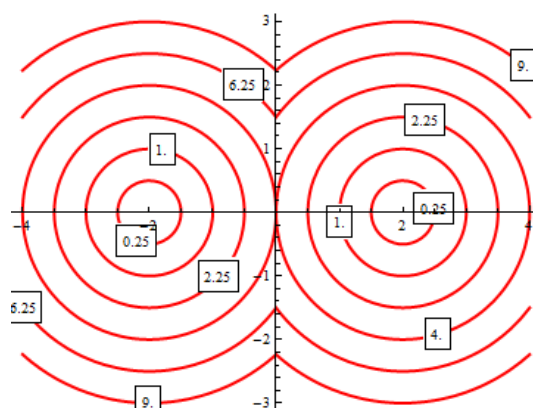
Au point  $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$  dans l'ensemble de niveau  $S_3$ , il apparaît que

$$T_{\begin{pmatrix} 2 \\ -1 \end{pmatrix}}^{en} S_c = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 \mid 2u = -v \right\}$$

et  $P_{\begin{pmatrix} 2 \\ -1 \end{pmatrix}}^{en} S_c = \left\{ \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} t \\ -2t \end{pmatrix} \right\}_{t \in \mathbb{R}} = \left\{ \begin{pmatrix} 2+t \\ -1-2t \end{pmatrix} \right\}_{t \in \mathbb{R}} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid 2x + y - 3 = 0 \right\}$  ♣

**Exemple 3.7.3.** Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  définie par  $F(x, y, z) = z^2 + (\sqrt{x^2 + y^2} - 2)^2$ , son gradient est

$$\nabla F(x, y, z) = \left( x \left( 1 - \frac{2}{\sqrt{x^2 + y^2}} \right), y \left( 1 - \frac{2}{\sqrt{x^2 + y^2}} \right), 2z \right) \quad \text{pourvu que } x^2 + y^2 \neq 0.$$

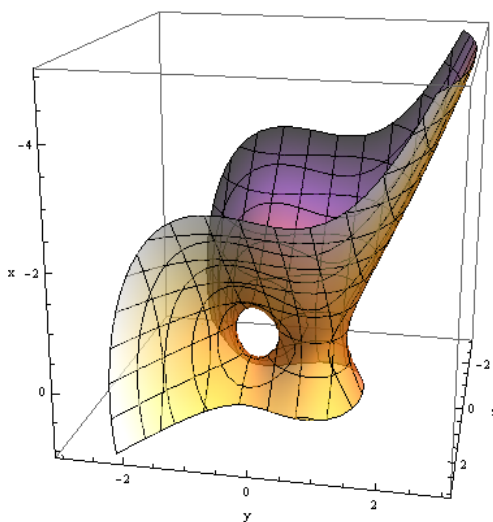
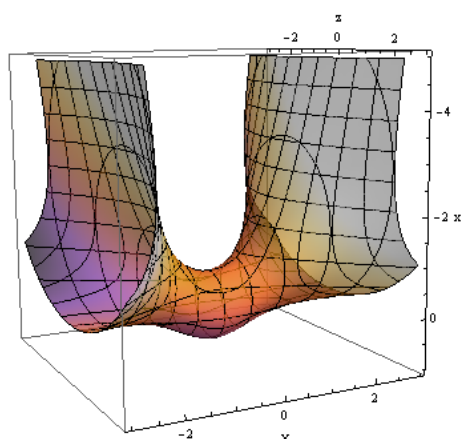


En particulier, au point  $(3, 4, 4)^T$  dans l'ensemble de niveau  $S_{25}$ ,  $\nabla F(3, 4, 4) = (\frac{9}{5}, \frac{12}{5}, 8)^T$ . Ceci donne  $T_{(3,4,4)}^{en} S_{25} = \{(u, v, w)^T \in \mathbb{R}^3 \mid 9u + 12v + 40w = 0\}$ ; une base de cet espace vectoriel est  $(4, -3, 0)^T$  et  $(0, 10, -3)^T$ . L'équation du plan tangent quant à elle est  $P_{(3,4,4)}^{en} S_{25} = \{(x, y, z)^T \in \mathbb{R}^3 \mid 9x + 12y + 40z - 235 = 0\}$ .

Dans le dessin ci-contre, n'est représentée qu'une partie de l'ensemble de niveau (c'est une tranche où  $y$  est fixé à 0)

Dans cet exemple, il est bon de souligner que le gradient n'est pas défini lorsque  $x = y = 0$ , et qu'il ne peut pas y être étendu par continuité. Ceci est dû au fait que les ensembles de niveau ( $S_c$  est l'ensemble des points à distance  $\sqrt{c}$  du cercle de rayon 2 contenu dans le plan  $xOy$ ) se comportent de manière "pointue" (*i.e.* qui ne s'approche pas bien linéairement) à cet endroit. Ceci paraît entre autres dans le dessin des coupes  $y = 0$  ci-après ( $y$  sont représentées les coupes de  $S_c$  où  $c = 1/4, 1, 9/4, 4, 25/4$  et 9). Petit exercice : où est l'ensemble de niveau  $S_0$  sur ce graphique ? ♣

**Exemple 3.7.4.** Finalement, voici un exemple d'ensemble de niveau pour une application. Soit  $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  l'application définie par  $\mathbf{F}(x, y, z) = \begin{pmatrix} z^2 + xy^2 + x^2 \\ z^2 + y(y^2 - 1) - x^2 \end{pmatrix}$ . Alors  $\nabla F_1(x, y, z) = (y^2 + 2x, 2xy, 2z)^T$  et  $\nabla F_2(x, y, z) = (-2x, 3y^2 - 1, 2z)^T$  (il y a beaucoup de points où ces deux gradients ne sont pas linéairement indépendants). Au point  $(\frac{\sqrt{3}}{4}, 0, \frac{\sqrt{5}}{4})^T$  dans l'ensemble de niveau  $S_{(1/2, 1/8)} = \mathbf{F}^{-1}(1/2, 1/8)$  l'espace tangent est formé des vecteurs  $\mathbf{v}$  satisfaisant  $\mathbf{v} \cdot (\frac{\sqrt{3}}{2}, 0, \frac{\sqrt{5}}{2})^T = 0$  et  $\mathbf{v} \cdot (-\frac{\sqrt{3}}{2}, 1, \frac{\sqrt{5}}{2})^T = 0$ . Autrement dit, c'est l'espace vectoriel engendré par  $(\frac{1}{\sqrt{3}}, 1, -\frac{1}{\sqrt{5}})^T$ . Ci-dessous les ensembles de niveau de  $F_1^{-1}(1/2)$  et  $F_2^{-1}(1/8)$ ; leur intersection n'est malheureusement pas simple à représenter.



♣

EXERCICE 44: Soit  $F(x, y) = x^2 + (|y| - 2)^2$  une fonction continue sur  $\mathbb{R}^2$ .

- a. Dessiner les ensembles de niveaux  $S_c = F^{-1}(c)$  pour  $c = 0, 1, 4$  et 9. [Indice : séparer les cas  $y \geq 0$  et  $y \leq 0$ .]

- b. Calculer le gradient de  $F$ .
- c. Identifier les points où le gradient n'est pas défini. Comment se manifeste géométriquement sur les ensembles de niveaux, l'impossibilité de définir le gradient ?
- d. Identifier les points où le gradient est nul. À quel[s] ensemble[s] de niveau[x] correspondent-ils ?
- e. Donner l'équation des plans tangents (en tant que sous-ensemble de  $\mathbb{R}^2$ ) en  $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$  et en  $\begin{pmatrix} \sqrt{2} \\ -2+\sqrt{2} \end{pmatrix}$  et les représenter graphiquement (soit sur le même dessin, soit sur un autre où la courbe de niveau est redessinée).

EXERCICE 45: Soit  $F$  fonction de classe  $C^1$  sur  $\mathbb{R}^2$  définie par  $F(x, y) = x(1 - x^2) + y^2$ .

- a. Faire une étude du polynôme  $p(x) = x^3 - x$ . En particulier, quelle est sa valeur maximale et minimale sur l'intervalle  $[-1, 1]$  ?
- b. Utiliser cette étude pour esquisser les ensembles de niveaux  $S_c = F^{-1}(c)$  pour  $c = -1, -\frac{2}{3\sqrt{3}}, 0, \frac{2}{3\sqrt{3}}$  et 1.
- c. Calculer le gradient de  $F$ . Identifier les points où il est nul. Comment se manifeste géométriquement sur les courbes de niveaux, l'annulation du gradient ?
- d. Donner l'équation des plans tangents (en tant que sous-ensemble de  $\mathbb{R}^2$ ) en  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  et en  $\begin{pmatrix} 2 \\ \sqrt{5} \end{pmatrix}$ . Les représenter graphiquement avec la courbe de niveau associée.

### 3.7.ii Espace des vecteurs vitesse

Avant de réinterpréter dans un sens non-linéaire la version EA2 ci-dessus, il est souhaitable de l'exprimer comme "un ensemble linéaire  $S$  est donné par un point de base et les points qu'il est possible d'atteindre depuis là par certaines droites". En fait, ceci est un écart (important) par rapport à la définition donnée en EA2. En effet, dit de cette façon rien n'oblige les droites à se trouver dans le même plan. Ainsi, cette description est celle (à translation près) d'un espace homogène et [centralement] symétrique (et non pas d'un espace affiné). Comme cette différence est importante, il vaut la peine de souligner un peu mieux ces concepts.

Un ensemble  $S \subset \mathbb{R}^m$  est dit homogène si  $\forall \lambda \in \mathbb{R}_{>0}, \mathbf{x} \in S \Rightarrow \lambda \mathbf{x} \in S$ . Un ensemble est dit centralement symétrique si  $\mathbf{x} \in S \Rightarrow -\mathbf{x} \in S$ .

Autrement dit, si  $S$  possède ces deux propriétés alors  $\forall \lambda \in \mathbb{R}, \mathbf{x} \in S \Rightarrow \lambda \mathbf{x} \in S$ .

Pour non-linéariser "la donnée d'un point et de droites qui passent par ce point", il suffit de remplacer "droite" par "courbe [paramétrée]".

Autrement dit, les ensembles dont une approximation sera ici tentée sont tout bonnement les ensembles connexes par arc ( $\forall \mathbf{x}, \mathbf{y} \in S, \exists \gamma: [-1, 1] \rightarrow S$  une courbe paramétrée continue telle que  $\gamma(-1) = \mathbf{x}$  et  $\gamma(1) = \mathbf{y}$ ). Avec une telle ambition (c'ad. une version qui s'applique dans autant de cas), la définition semblera évidemment souple (c'ad. moins accorte), mais elle est aussi très utile.

Ainsi, soit  $S$  un ensemble connexe par arc et  $\mathbf{x}$  un point de base où un linéarisé de  $S$  est à construire. Dans ce contexte, il y a des courbes paramétrées qui relient  $\mathbf{x}$  à d'autres points de  $S$  (sauf si  $S = \{\mathbf{x}\}$ ). En mettant bout à bout de telles courbes, l'existence de courbe paramétrée  $\gamma$



passant par  $\mathbf{x}$  (i.e. telle que  $\text{Im } \gamma \ni \mathbf{x}$ ) se démontre sans peine. Lorsqu'elles sont dérivables en  $\mathbf{x}$  ces courbes s'approchent par des droites, ce qui permet de tenter une approximation de  $S$  en  $\mathbf{x}$ . Ces considérations motivent la définition suivante.

**Définition 3.7.5.** Soit  $S \subset \mathbb{R}^m$  un sous-ensemble connexe par arc. Alors l'espace tangent des vitesses<sup>1</sup> à  $S$  en  $\mathbf{x} \in S$ , noté  $T_{\mathbf{x}}^{\text{vit}} S$  est

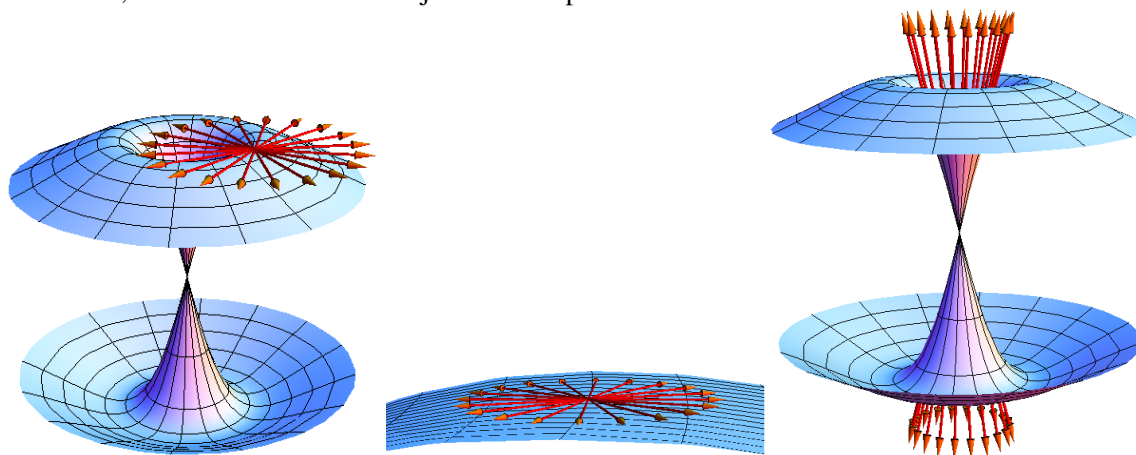
$$T_{\mathbf{x}}^{\text{vit}} S = \{\mathbf{v} \in \mathbb{R}^m \mid \exists \gamma: [-1, 1] \rightarrow S \text{ de classe } C^1 \text{ telle que } \gamma(0) = \mathbf{x} \text{ et } \gamma'(0) = \mathbf{v}\}.$$

En plus de mots : c'est l'ensemble des vecteurs  $\mathbf{v}$  tels qu'il existe une courbe paramétrée  $\gamma: [-1, 1] \rightarrow S$  de classe  $C^1$  satisfaisant  $\gamma(0) = \mathbf{x}$  et  $\gamma'(0) = D_1 \gamma(0) = \mathbf{v}$ .

Comme plus haut, le plan tangent est  $P_{\mathbf{x}}^{\text{vit}} S = \{\mathbf{y} \in \mathbb{R}^m \mid \exists \mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S, \mathbf{y} = \mathbf{x} + \mathbf{v}\}$ . ★

Le prix de tant de flexibilité (la restriction aux ensembles connexes par arc est sans importance) est que cet espace n'est, *a priori*, PAS un espace vectoriel, c'àd. si  $\mathbf{v}, \mathbf{w} \in T_{\mathbf{x}}^{\text{vit}} S$ , alors  $\mathbf{v} + \mathbf{w}$  n'appartient pas nécessairement à  $T_{\mathbf{x}}^{\text{vit}} S$ . Il est, par contre, possible de montrer qu'il est homogène, c'àd. que si  $\mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S$  alors  $\forall a \in \mathbb{R}, a\mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S$ .

Côté représentation graphique, voici ci-dessous trois dessins d'un sous-ensemble de  $\mathbb{R}^3$ . Le premier dessin (à gauche) représente une vingtaine de vecteurs vitesse en un point. En regardant de plus près (deuxième dessin, au milieu), il transpire que ces vecteurs se trouvent tous dans un même plan ;  $T^{\text{vit}}$  forme un espace vectoriel en ce point. Par contre (troisième dessin, à droite), au point où il y a un étranglement, les vecteurs vitesses ne forment définitivement pas un espace vectoriel ; ils se répartissent en fait dans un cône. Noter qu'ils demeurent une très bonne approximation linéaire de l'ensemble, même si celle-ci ne sera jamais un espace vectoriel.



**Exemple 3.7.6.** Soit  $S \subset \mathbb{R}^2$  l'ensemble donné par l'union du cercle de rayon 1 centré en  $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$  et du cercle de rayon centré en  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . Alors  $T^{\text{vit}} S$  est, aux points  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  et  $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ , l'espace vectoriel engendré par le vecteur  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , tandis qu'aux points  $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$  et  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  c'est l'espace engendré par le vecteur  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . ♣

1. Certaines définitions (encore plus générale) demandent que  $\gamma$  soit seulement définie sur  $[0, 1]$  et il faut alors considérer la vitesse en 0 en prenant la limite à gauche seulement  $\gamma'(0^+) := \lim_{t \rightarrow 0^+} \frac{\gamma(t) - \gamma(0)}{t}$ .

EXERCICE 46: Soit  $S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid xy = 0 \right\}$ .

- Réaliser  $S$  comme  $S_0 = F^{-1}(0)$  d'une fonction (de classe  $C^1$ )  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ .
- Montrer que  $T_0^{vit} S \subsetneq T_0^{ev} S_0$ .
- Trouver une application  $G : \mathbb{R} \rightarrow \mathbb{R}^2$  de classe  $C^1$  de  $S$  tel que  $G(0) = \mathbf{0}$  et  $T_0^{par} G \subsetneq T_0^{vit} S$ .

Comme il sera vu aux exercices, cette définition est parfois la meilleure approximation linéaire d'un ensemble. Au chapitre 4, il sera aussi démontré que certains ensembles  $S$  qui sont l'image de courbes paramétrées  $C^\infty$  (et non-triviale) ont un  $T^{vit}$  qui se réduit à  $\{\mathbf{0}\}$  en certains points.

EXERCICE 47: Soit  $F : (\mathbb{R} \setminus \{0\}) \times \mathbb{R} \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = e^{2 \frac{1+y^2}{3+y^2} \ln|x|}$ .

- Trouver la fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  telle que  $F(x, y) = |x|^{f(y)}$ . Faire une étude de  $f$  et la représenter graphiquement, en particulier, marquer les points où  $f(y) = 1$ .
- $F(x, y)$  peut-elle s'étendre en  $x = 0$  ?
- Esquisser  $\Gamma_F = \{(x, y, z) \in \mathbb{R}^3 \mid z = F(x, y)\}$  le graphe de  $F$ . Il peut-être commode de tracer le graphe de  $F(x, 0)$ ,  $F(x, \pm 1)$  et  $F(x, \pm 2)$  (des fonctions de  $\mathbb{R} \rightarrow \mathbb{R}$ ).
- Calculer  $D_1 F$  et  $D_2 F$ . Montrer que  $D_2 F$  est continue sur  $\mathbb{R}^2$ . Montrer que  $D_1 F$  est continue sur  $\mathbb{R}^2 \setminus \{0\} \times [-1, 1]$ . [Indice :  $\forall \alpha > 0, \lim_{x \rightarrow 0} |x|^\alpha \ln|x| = 0$ .]
- Ainsi, le gradient de  $F$  ne peut être défini sur  $\{0\} \times [-1, 1]$ . Comment se manifeste cela de manière géométrique sur le graphe ?
- Donner le plan tangent au graphe en un point de votre choix.

### 3.7.iii Image d'une application linéarisée

Si la version EA2 a subi un traitement légèrement alambiqué, l'adaptation de la version EA3 est, elle, d'une simplicité évidente. Au lieu de regarder  $S$  comme l'image d'une application affine, notre ensemble  $S$  à linéariser sera l'image d'une application (pas nécessairement affine).

Ainsi, soit  $V \subset \mathbb{R}^\ell$  et  $\mathbf{G} : V \rightarrow \mathbb{R}^m$  une application (pour l'instant seulement  $C^0$ ). L'objet de notre étude est ici  $S = \text{Im } \mathbf{G}$ . À un point  $\mathbf{x} \in V$  sera associé comme ensemble linéarisé l'image de la dérivée de  $\mathbf{G}$  (l'approximation linéaire de  $\mathbf{G}$ ). Il y a un problème dans cette terminologie :  $\mathbf{x}$  est dans l'image de  $\mathbf{G}$  et  $\mathbf{G}$  n'est pas nécessairement injective. Pour éviter ce problème (et tenter d'alléger les notations de fonctions inverses encombrantes), l'espace tangent ici sera pris par rapport à  $\mathbf{y} \in V$ .

**Définition 3.7.7.** Un paramétrage d'un ensemble  $S \subset \mathbb{R}^m$  est une application  $\mathbf{G} : V \rightarrow \mathbb{R}^m$  de classe  $C^1$  telle que  $\text{Im } \mathbf{G} \subset S$ .

Soit  $V \subset \mathbb{R}^\ell$  un ouvert et  $\mathbf{G} : V \rightarrow \mathbb{R}^m$  une application dérivable en  $\mathbf{y} \in V$ . Alors l'espace tangent en tant que paramétrage de  $S' := \text{Im } \mathbf{G}$  en  $\mathbf{y}$  est défini par

$$T_{\mathbf{y}}^{par} S' := T_{\mathbf{y}}^{par} \mathbf{G} := \text{Im } \mathbf{G}_*(\mathbf{x}).$$

Le plan tangent de ce paramétrage en  $\mathbf{y}$  est  $P_{\mathbf{y}}^{par} \mathbf{G} = \{\mathbf{y} \in \mathbb{R}^m \mid \exists \mathbf{v} \in T_{\mathbf{x}}^{vit} S, \mathbf{y} = \mathbf{x} + \mathbf{v}\}$ . ★



**Attention :** il n'est pas requis que le paramétrage soit surjectif !

L'écriture concise ci-dessus peut aussi s'écrire sous la forme plus explicite suivante :

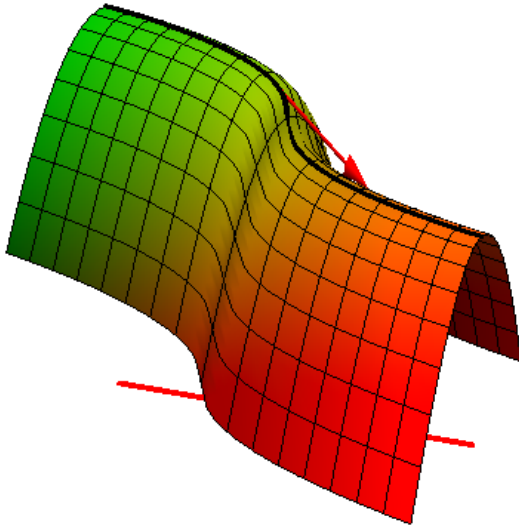
$$T_{\mathbf{y}}^{par} \mathbf{G} = \left\{ \mathbf{v} \in \mathbb{R}^m \mid \exists c_1, c_2, \dots, c_\ell, \text{ tels que } \mathbf{v} = \sum_{i \in \bar{\ell}} c_i D_i \mathbf{G}(\mathbf{y}) \right\}$$

En effet, il est alors évident que l'espace vectoriel obtenu est de dimension 0 à  $\ell$ . Cet espace tangent dépend du paramétrage, *i.e.* pour une image  $S'$  fixée deux paramétrages peuvent avoir un plan tangent très différent.

Pour ceux qui ne serait satisfait de la justification ci-haut voici une interprétation (qui n'est pas rigoureuse). S'il est admis que  $\mathbf{G} : V \rightarrow S \subset \mathbb{R}^m$  est un bon paramétrage de  $S$ , alors une courbe  $\gamma : [a, b] \rightarrow S$  devrait pouvoir être exprimée comme  $\mathbf{G} \circ \phi$  où  $\phi : [a, b] \rightarrow V$  est une courbe. Or par la règle de dérivation en chaîne,  $\gamma'(t) = \mathbf{G}_*(\phi(t))\phi'(t)$  ainsi  $\gamma'(t) \in \text{Im } \mathbf{G}_*(\phi(t))$ . Il est alors assez naturel de tenter d'approcher l'espace tangent des vecteurs vitesse par cet espace tangent des paramétrages.

Un exemple particulièrement important de ce concept est l'idée de l'espace (ou du plan) tangent au graphe d'une fonction.

Soit  $U \subset \mathbb{R}^m$  un ouvert,  $F : U \rightarrow \mathbb{R}$  une fonction dérivable sur  $U$ . Un élément de  $\mathbb{R}^m \times \mathbb{R}$  s'écrit  $\begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} = (x_1, x_2, \dots, x_m, z)^\top$ . Soit donc  $\Gamma_F = \left\{ \begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R} \mid z = F(\mathbf{x}) \right\}$  le graphe de  $F$ . Le plan tangent à  $\begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} \in \Gamma_F$  est engendré par les vecteurs (basés en  $\begin{pmatrix} \mathbf{x} \\ z \end{pmatrix}$ )  $(D_i F(\mathbf{x}))^\top$  où  $i \in \bar{m}$ . En effet, dans la direction  $\mathbf{e}_i$  la variation infinitésimale de  $F$  est  $D_i F(\mathbf{x})$ .



Sur le dessin ci-contre est représenté un graphe de fonction  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ . En bas en rouge se trouve une droite parallèle à l'axe des  $x$  et sur le graphe est tracé en noir son image. La flèche sur le graphe est le vecteur  $(1, 0, D_1 F(x, y))^\top$  basé au point  $(x, y, F(x, y))$ . S'il était possible de regarder très près, il apparaîtrait que cette flèche est bien la tangente à la courbe tracée en noir.

En particulier pour  $m = 2$ , au point  $(x, y, F(x, y))^\top$  du graphe  $\Gamma_F = \{(x, y, z)^\top \in \mathbb{R}^3 \mid z = F(x, y)\}$  le plan tangent est engendré par  $(1, 0, D_1 F(x, y))^\top$  et  $(0, 1, D_2 F(x, y))^\top$ . Le produit vectoriel permet aussi de trouver le vecteur normal (unique à un scalaire près) qui définit ce plan (espace vectoriel de dimension 2 dans  $\mathbb{R}^3$ ) :

$$\begin{pmatrix} 1 \\ 0 \\ D_1 F(x, y) \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ D_2 F(x, y) \end{pmatrix} = \begin{vmatrix} \mathbf{e}_1 & 1 & 0 \\ \mathbf{e}_2 & 0 & 1 \\ \mathbf{e}_3 & D_1 F(x, y) & D_2 F(x, y) \end{vmatrix} = \begin{pmatrix} -D_1 F(x, y) \\ -D_2 F(x, y) \\ 1 \end{pmatrix}$$

**Exemple 3.7.8.** Soit de nouveau  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = x^2 - y^2$ . Au point  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  l'espace tangent au graphe est donné par l'orthogonalité au vecteur (basé en  $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ )  $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$ . L'équation du plan est ainsi  $-2x + 2y + z = 0$ . ♣

En général, pour  $m \geq 3$ , il est possible de montrer que le vecteur normal est proportionnel à

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ D_1 F(x, y) \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ D_2 F(x, y) \end{pmatrix} \wedge \dots \wedge \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ D_m F(x, y) \end{pmatrix} = (-1)^m \begin{pmatrix} -\nabla F(\mathbf{x}) \\ 1 \end{pmatrix}.$$

EXERCICE 48: Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = x^2(1 - x^2) + (y^2 - 1)y$ . Soit  $\Gamma_F = \{(x, y, z) \in \mathbb{R}^3 \mid z = F(x, y)\}$  le graphe de  $F$ .

- Calculer  $D_1 F$  et  $D_2 F$ .
- $F$  donne lieu au paramétrage habituel de son graphe  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  défini par  $\Phi(x, y) = (x, y, F(x, y))^T$ . Donner le plan tangent  $\mathbf{P}_{(x, y)}^{par} \Phi$  en un point  $\begin{pmatrix} x \\ y \end{pmatrix}$  arbitraire. Quelle est la dimension de cet espace ?
- $\Gamma_F$  a peut aussi être vu comme l'ensemble de niveau  $S_0 := G^{-1}(0)$  de la fonction  $G; \mathbb{R}^3 \rightarrow \mathbb{R}$  définie par  $G(x, y, z) = z - F(x, y)$ . Calculer le gradient de  $G$ .
- Un calcul du plan tangent  $\mathbf{P}^{en} S_0$  devrait donner le même résultat qu'en b. Montrer seulement que  $(\nabla G)(\Phi(x, y))$  est orthogonal à  $D_1 \Phi(x, y)$  et  $D_2 \Phi(x, y)$ .

### 3.7.iv Premier lien entre les trois définitions

La dérivation composée permet de donner un premier lien entre ces trois concepts. En effet, si  $U \subset \mathbb{R}^m$ ,  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application  $\mathcal{C}^1$ ,  $V \subset \mathbb{R}^\ell$  et  $\mathbf{G} : V \rightarrow \mathbb{R}^m$  sont tels que  $\text{Im } \mathbf{G} \subset U$  et  $\forall \mathbf{y} \in V, \mathbf{F} \circ \mathbf{G}(\mathbf{y}) = \mathbf{c}$  (i.e.  $\mathbf{c}$ 'est une application constante), alors trois définitions sont à disposition pour décrire un seul ensemble  $S : S = \mathbf{F}^{-1}(\mathbf{c})$  est un ensemble de niveau,  $\mathbf{G}$  est un paramétrage de  $S$ , et  $S$  est un ensemble où des courbes paramétrées peuvent se tracer.

Fort heureusement, les trois versions du plan tangent forment dans ces cas une inclusion simple.

#### Proposition 3.7.9

Soit  $S \subset \mathbb{R}^m$ .

Si  $\mathbf{G}$  est un paramétrage de classe  $\mathcal{C}^1$  de  $S$ , et  $\mathbf{x} = \mathbf{G}(\mathbf{y})$  alors  $\mathbf{T}_{\mathbf{y}}^{par} \mathbf{G} \subset \mathbf{T}_{\mathbf{x}}^{vit} S$ .

Si  $\mathbf{F}$  est une application de classe  $\mathcal{C}^1$  telle que  $S_{\mathbf{c}} = \mathbf{F}^{-1}(\mathbf{c}) \supset S$  alors si  $\mathbf{x} \in S$ ,  $\mathbf{T}_{\mathbf{x}}^{vit} S \subset \mathbf{T}_{\mathbf{x}}^{en} S_{\mathbf{c}}$

Avant d'entamer la démonstration, il est bon de se rappeler que, si  $\gamma$  est une courbe paramétrée (de classe  $\mathcal{C}^1$ ) alors  $\gamma'(t) = \gamma_*(t)$ .

DÉMONSTRATION: Pour la première de ces deux affirmations, il suffit de constater que si  $\mathbf{v} \in \mathbf{T}_{\mathbf{y}}^{par} \mathbf{G}$  alors  $\mathbf{v} = \sum_{i \in \bar{I}} a_i D_i \mathbf{G}(\mathbf{y})$  où  $a_i \in \mathbb{R}$ . Autrement dit,  $\mathbf{v} = \mathbf{G}_*(\mathbf{y})\mathbf{a}$  où  $\mathbf{a} = (a_1, a_2, \dots, a_\ell)^T \in \mathbb{R}^\ell$ . Soit  $\sigma : [-1, 1] \rightarrow \mathbb{R}^\ell$  la courbe paramétrée de classe  $\mathcal{C}^1$  définie par  $\sigma(t) = \mathbf{y} + t\mathbf{a}$ . Soit  $\gamma = \mathbf{G} \circ \sigma$ . Alors,  $\gamma$  est de classe  $\mathcal{C}^1$ ,  $\text{Im } \gamma \subset \text{Im } \mathbf{G} \subset S$ ,  $\gamma(0) = \mathbf{x}$  et, par le théorème sur la dérivation en chaîne 3.4.1,  $\gamma'(0) = \mathbf{G}_*(\sigma(0))\sigma'(0) = \mathbf{G}_*(\mathbf{y})\mathbf{a} = \mathbf{v}$ . Ainsi  $\mathbf{v} \in \mathbf{T}_{\mathbf{x}}^{vit} S$ .

La seconde assertion est la reprise de l'argument sur l'orthogonalité du gradient aux ensembles de niveaux. En effet, soit  $\mathbf{v} \in T_{\mathbf{x}}^{vit} S$  et soit  $\gamma$  de classe  $C^1$  telle que  $\text{Im } \gamma \subset S$ ,  $\gamma(0) = \mathbf{x}$  et  $\gamma'(0) = \mathbf{v}$ . Alors  $\mathbf{F} \circ \gamma(t) = \mathbf{c}$  car  $\text{Im } \gamma \subset S \subset S_{\mathbf{c}} = \mathbf{F}^{-1}(\mathbf{c})$ . En particulier, par la règle de dérivation en chaîne 3.4.1,  $\mathbf{F}_*(\gamma'(0)) = \mathbf{0}$  d'où  $\mathbf{F}_*(\mathbf{x})\mathbf{v} = \mathbf{0}$ . Or la  $i^{\text{ème}}$  ligne ( $i \in \bar{n}$ ) de  $\mathbf{F}_*(\mathbf{x})$  est  $\nabla F_i(\mathbf{x})$ , d'où  $\forall i \in \bar{n}$ ,  $\nabla F_i(\mathbf{x}) \cdot \mathbf{v} = 0$  ou, par définition,  $\mathbf{v} \in T_{\mathbf{x}}^{en} S_{\mathbf{c}}$ . ■

Quelle que soit la manière dont un sous-ensemble de  $\mathbb{R}^m$  est défini (*i.e.* comme ensemble de niveau dans ET1, comme image d'une application dans ET3, ou tout simplement comme sous-ensemble dans ET2), un point est singulier lorsque les trois définitions de plan tangent ne peuvent coïncider (ou ne peuvent être définies). Il faudra attendre le théorème des fonctions implicites 6.2.5 pour avoir une (très) bonne caractérisation des points réguliers (*c'ad.* qui ne sont pas singuliers, *c'ad.* où les trois notions d'espaces tangents coïncident).

Tentons de motiver une démonstration en se ramenant à de l'algèbre linéaire. Supposons que  $A$  est une matrice  $\ell \times m$  (*i.e.* une application linéaire de  $\mathbb{R}^{\ell}$  dans  $\mathbb{R}^m$ ) et que, de plus,  $A$  est injective en tant qu'application linéaire (*i.e.*  $\text{Ker } A = \{0\}$ ). Étant donné une base de l'image  $\text{Im } A$  (et de son complémentaire), il est aisé de construire une matrice  $B$  de taille  $m \times n$  telle que  $\text{Ker } B = \text{Im } A$  (et  $n = m - \dim \text{Im } A = m - \ell$ ) et  $B$  est surjective en tant qu'application linéaire. Autrement dit,  $B$  est un isomorphisme du complémentaire de  $\text{Im } A$  avec  $\mathbb{R}^n$ .

Inversement, étant donné  $B$  de taille  $m \times n$  qui est surjective en tant qu'application linéaire, trouver une base de  $\text{Ker } B$  permet très facilement de construire une matrice  $A$  de taille  $\ell \times m$  telle que  $\text{Im } A = \text{Ker } B$  (et  $\ell = \dim \text{Ker } B = m - n$ ).

Autrement dit, dans les deux cas, si une des applications satisfait certaines propriétés, il est possible de reconstruire la seconde. Dans les deux cas  $A$  décrit l'ensemble  $\text{Im } A = S$ , tandis que  $B$  décrit le quotient de l'espace par cet ensemble (autrement les différentes tranches de  $\mathbb{R}^m/S$ ). Il est possible de décomposer  $\mathbb{R}^m$  en deux sous-espaces complémentaires :  $S = \text{Im } A = \text{Ker } B$  qui est isomorphe via  $A$  à  $\mathbb{R}^{\ell}$  et son complémentaire qui est isomorphe via  $B$  à  $\mathbb{R}^n$ .

En revenant aux applications non-linéaires (dont l'image de l'une est contenue dans l'ensemble de niveau de l'autre), demander que  $A$  est un isomorphisme sur son image (*i.e.* est injective) revient à demander que le paramétrage  $\mathbf{G}$  soit (à peu de chose près) une bonne description de son image (formellement,  $\mathbf{G}_*(\mathbf{y})$  est injectif si les  $D_i \mathbf{G}(\mathbf{y})$  sont linéairement indépendants). Demander que  $B$  soit un isomorphisme à noyau près (*i.e.* est surjective) revient à demander que l'application  $\mathbf{F}$  a la propriété que, en variant  $\mathbf{c}$ , les ensembles de niveau  $S_{\mathbf{c}}$  de  $\mathbf{F}$  "s'empilent" pour bien recouvrir un voisinage de l'ensemble de niveau original (formellement,  $\mathbf{F}_*(\mathbf{x})$  est surjective si  $n \leq m$  et, pour  $i \in \bar{n}$ , les  $\nabla F_i(\mathbf{x})$  sont linéairement indépendants). Qui plus est, dans ces deux situations ( $A$  injective et  $B$  surjective), la dimension de l'espace tangent est extrême ( $T^{\text{par}}$  est de dimension maximale et  $T^{\text{en}}$  de dimension minimale). En regardant de plus près, elles sont même égales.

Quelques part, la proposition 3.7.9 ne permet pas de retrouver tout ce qui se passe dans le cas linéaire. Mais elle permet quand même d'en retrouver un peu :

#### Corollaire 3.7.10

Soit  $U \subset \mathbb{R}^m$  ouvert, soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application de classe  $C^1$  et soit  $S_{\mathbf{c}} = \mathbf{F}^{-1}(\mathbf{c})$ . Soit

### 3.7 - Espace(s) tangent(s)

$V \subset \mathbb{R}^\ell$  un ouvert, soit  $\mathbf{G} : V \rightarrow \mathbb{R}^m$  un paramétrage de classe  $C^1$  de  $S$ , et  $\mathbf{x} = \mathbf{G}(\mathbf{y})$ .

1. Si les  $D_i \mathbf{G}(\mathbf{y})$  (pour  $i \in \overline{\ell}$ ) sont linéairement indépendants, alors  $\dim T_{\mathbf{y}}^{par} \mathbf{G} = \ell$ .
2. Si les  $\nabla F_j(\mathbf{x})$  (pour  $j \in \overline{n}$ ) sont linéairement indépendants, alors  $\dim T_{\mathbf{x}}^{en} S_c = m - n$ .
3. Si  $\dim T_{\mathbf{y}}^{par} \mathbf{G} = \dim T_{\mathbf{x}}^{en}$  alors  $T_{\mathbf{x}}^{vit} S_c$  est un espace vectoriel identique aux deux autres.

DÉMONSTRATION: Les deux premières affirmations sont de simples calculs sur les espaces vectoriels. Pour la troisième, il suffit de voir que  $T^{par}$  et  $T^{en}$  sont deux espaces vectoriels de même dimension dont le premier est inclus dans le second. Par conséquent, ils doivent être égaux. Mais comme  $T^{vit}$  est coincé entre  $T^{par}$  et  $T^{en}$ , il est lui aussi égal aux deux autres et est donc aussi un espace vectoriel. ■

En fait, retrouver la situation idéale des applications linéaire demande beaucoup de travail : c'est essentiellement le contenu du théorème des applications implicites, et sa démonstration demande est loin d'être triviale.

EXERCICE 49: Soit  $S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid xy = 0 \right\}$ .

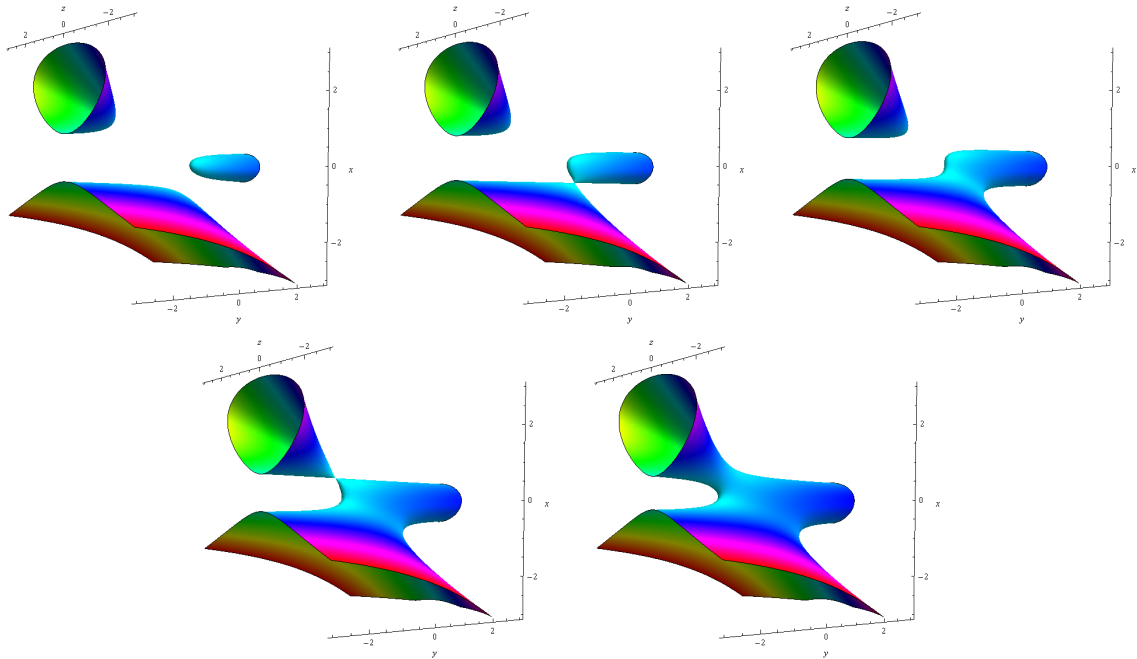
- a. Réaliser  $S$  comme l'ensemble de niveau  $S_0 = F^{-1}(0)$  d'une fonction  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  de classe  $C^1$ .
- b. Montrer que  $T_{\mathbf{0}}^{vit} S \subsetneq T_{\mathbf{0}}^{en} S_0$ .
- c. Trouver un paramétrage  $G : \mathbb{R} \rightarrow S$  de classe  $C^1$  de  $S$  tel que  $G(0) = \mathbf{0}$  et  $\{0\} \subsetneq T_0^{par} G \subsetneq T_0^{vit} S$ .

EXERCICE 50: Soit  $S = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 = z^2 \right\}$  le cône (double).

- a. Réaliser  $S$  comme  $S_0 = F^{-1}(0)$  d'une fonction (de classe  $C^1$ )  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$ .
- b. Déterminer  $T^{en} S_0$  en chaque point. Quel est sa dimension ?
- c. Déterminer  $T_0^{vit} S$ . [Utiliser que  $S$  est fermé et homogène :  $\mathbf{x} \in S \Rightarrow \forall \lambda \in \mathbb{R}, \lambda \mathbf{x} \in S$ .]
- d. Trouver une application  $\mathbf{G} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  de classe  $C^1$  de  $S$  tel que  $\mathbf{G}(\mathbf{0}) = \mathbf{0}$ . Quel est la dimension de  $T^{par} \mathbf{G}$  en fonction du point ?
- e. Montrer que  $T_0^{par} \mathbf{G} \subsetneq T_0^{vit} S \subsetneq T^{en}$ . Pourquoi n'était-il ni nécessaire de calculer  $T^{par}$  ni  $T^{en}$  pour le savoir ?

EXERCICE 51: Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y, z) = x^3 + z^2 + x^2 y - \frac{y}{4}$ .

- a. Calculer le gradient de  $F$ , et dire en quels points il s'annule.
- b. Ci-dessous sont dessinés 5 ensembles de niveau. En deux valeurs, l'ensemble de niveau présente une singularité (c'ad. que le plan tangent ne peut être défini par orthogonalité au gradient, e.g. quand il est nul), de quel  $S_c$  s'agit-il ?



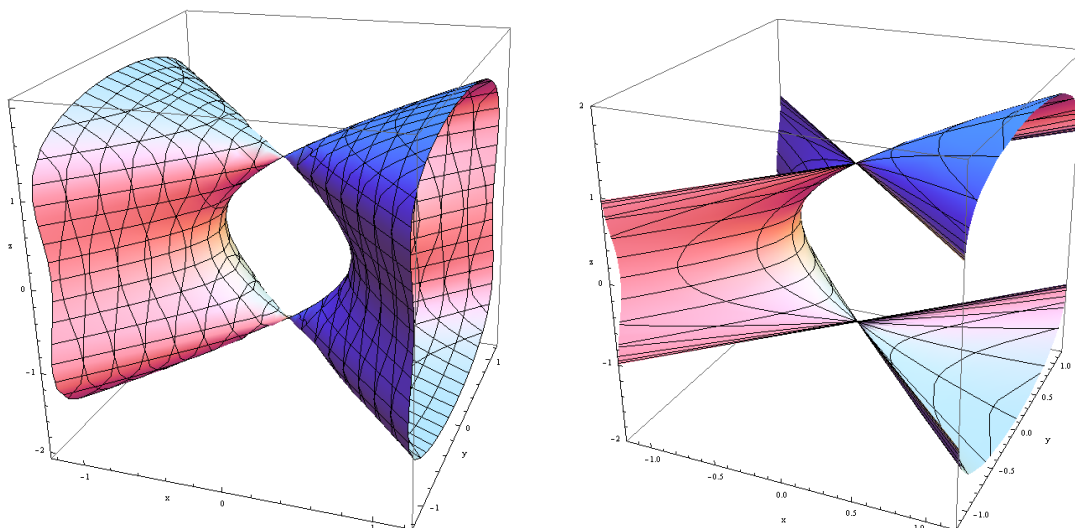
EXERCICE 52: Il est possible de décrire l'ellipse de plusieurs façons. Soient  $a, b \in \mathbb{R}_{>0}$  les deux demi-axes de l'ellipse.

- La courbe paramétrée  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} a \sin t \\ b \cos t \end{pmatrix}$  a pour image cette ellipse (centrée en  $\mathbf{0}$  et les axes le long des axes de coordonnées). Calculer le vecteur vitesse  $\gamma'(t)$ , la matrice de Jacobi  $\gamma_*(t)$  et la dérivée en  $t$ . [Il s'agit à peu de chose près des mêmes objets...]
- En déduire, pour tout  $s \in \mathbb{R}$  l'espace tangent  $T_s^{\text{par}} \gamma$ . De quel dimension est-il ?
- L'ellipse est aussi la solution de l'équation  $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ . Autrement dit, c'est un ensemble de niveau de la fonction  $F: \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par  $F(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ . Calculer le gradient et la dérivée de  $F$  en  $(x, y)$ . [Il s'agit à peu de chose près des mêmes objets...]
- De quel ensemble de niveau s'agit-il ? Décrire l'espace tangent d'un point  $\begin{pmatrix} x \\ y \end{pmatrix}$  dans cet ensemble de niveau. De quel dimension est-il ?
- Que pouvez-vous en déduire sur l'espace tangent des vecteurs vitesses ?

EXERCICE 53: Soit  $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 - y^2 = (z^2 - 1)^2\}$ . Soit  $\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  l'application définie par  $\Phi(r, t) = \begin{pmatrix} (r^2-1) \cos t \\ (r^2-1) \sin t \\ r \end{pmatrix}$ . L'image de  $\Phi$  est contenue dans  $S$ . Ci-dessous à gauche se trouve une représentation graphique de  $S$ . Ci-dessous à droite se trouve une représentation graphique de l'image de  $\Phi$ .

- Calculer les dérivées partielles de  $\Phi$  et donner une base de  $T_{(r,t)}^{\text{par}} \Phi$ . De quelle dimension est cet espace vectoriel ?
- Décrire  $S$  comme l'ensemble de niveau  $S_0 = F^{-1}(0)$  d'une fonction  $F: \mathbb{R}^3 \rightarrow \mathbb{R}$  et donner le gradient de  $F$ .
- Quelles conditions doivent satisfaire les vecteurs de  $T_{(x,y,z)}^{\text{en}} S_0$  ? De quelle dimension est cet espace vectoriel ?

- d. En déduire la dimension de  $T_{(x,y,z)}^{vit} S$  pour tous les points sauf deux dans l'image de  $\Phi$ . Quels sont ces deux points ?



L'exercice suivant est une version générale de l'exercice 48 ; il montre que pour le graphe d'une fonction  $C^1$ , il y a des notions de plan tangents qui coïncident toutes bien (pour un choix assez naturel de paramétrage et d'ensemble de niveau).

EXERCICE 54: Soit  $G : \mathbb{R}^\ell \rightarrow \mathbb{R}^m$  une application de  $\mathbb{R}^\ell$  dans  $\mathbb{R}^m$  de classe  $C^1$ .

- L'application  $F : \mathbb{R}^\ell \rightarrow \mathbb{R}^{\ell+m}$  définie par  $F(x) = \begin{pmatrix} x \\ G(x) \end{pmatrix}$  prend valeur dans le graphe de  $G$ . Montrer que  $F$  est injective.
- Montrer que  $F_*(x)$  est toujours injective (en tant qu'application linéaire). En déduire que  $\dim T_x^{par} F = \ell$ .
- Soit  $H : \mathbb{R}^{\ell+m} \rightarrow \mathbb{R}^m$  définie par  $H(x, y) = G(x) - y$  où  $x \in \mathbb{R}^\ell$  et  $y \in \mathbb{R}^m$ . Décrire  $H_*(x)$  ; en conclure que  $H_*(x)$  est surjective, puis que  $\text{Ker } H_*(x)$  est de dimension  $\ell$ .
- Montrer que  $\Gamma_G = H^{-1}(0)$ . Dire pourquoi l'espace tangent des vecteur vitesse est un espace vectoriel de dimension  $\ell$ .

### 3.A Dérivée de Gâteaux

Soit  $U \subset \mathbb{R}^m$  un ouvert, et  $F : U \rightarrow \mathbb{R}^n$ . La dérivée de Gâteaux (ou dérivée directionnelle) de  $F$  de  $F$  en  $x \in U$  dans la direction  $y \in \mathbb{R}^m$  (où  $\|y\|$  ne vaut pas nécessairement 1), notée  $d_G F(x; y)$ , est définie par

$$d_G F(x; y) = D_y F(x) = \lim_{t \rightarrow 0} \frac{F(x + ty) - F(x)}{t} = \left. \frac{d}{dt} F(x + ty) \right|_{t=0}$$

si la limite existe. Lorsque cette limite existe pour tout  $y \in \mathbb{R}^n$ ,  $F$  est dite dérivable au sens de Gâteaux en  $x$ . De la même façon que la dérivée usuelle (ou de Fréchet) associe à un point où  $F$  est dérivable une application (linéaire), en chaque point  $x \in U$  où  $F$  est dérivable au sens de Gâteaux il y a une application (la dérivée de Gâteaux)

$$d_G F(x; \cdot) : \mathbb{R}^m \rightarrow \mathbb{R}^n.$$



Cette application n'est par contre pas nécessairement linéaire. Elle n'est *a priori* que homogène, *i.e.* pour tout scalaire  $t \in \mathbb{R}$  et tout  $\mathbf{y} \in \mathbb{R}^n$ ,  $d_G \mathbf{F}(\mathbf{x}; t\mathbf{y}) = t\mathbf{F}(\mathbf{x}; \mathbf{y})$ .

**Exemple 3.A.1.** Soit  $F$  la fonction (homogène de degré 1)  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par

$$F(x, y) = \begin{cases} \frac{x^3}{x^2+y^2} & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 0 & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{cases}$$

$F$  est continue sur  $\mathbb{R}^2$  et elle est dérivable au sens de Fréchet en tout point sauf  $\mathbf{0}$ . De plus,  $F$  est dérivable au sens de Gâteaux en tout point,  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  compris ; sa dérivée de Gâteaux en  $\mathbf{0}$  est

$$d_GF(0, 0; a, b) = F(a, b).$$

Ainsi  $d_GF(\mathbf{0}; \cdot)$  est continue, homogène, mais n'est pas une application linéaire. ♣

Lorsqu'une application  $\mathbf{F}$  a une dérivée (de Fréchet)  $d\mathbf{F}_{\mathbf{x}}$  en un point  $\mathbf{x}$ , sa dérivée de Gâteaux en ce point existe et  $d_G \mathbf{F}(\mathbf{x}; \mathbf{y}) = d\mathbf{F}_{\mathbf{x}}(\mathbf{y})$ . Par contre, comme l'exemple ci-dessus le montre, l'existence d'une dérivée de Gâteaux est insuffisante pour être dérivable au sens de Fréchet. De plus, la fonction  $F$  de l'exemple 2.4.4 possède une dérivée de Gâteaux en  $\mathbf{0}$  (encore une fois  $d_GF(\mathbf{0}; \mathbf{y}) = F(\mathbf{y})$ ), sans même être continue en  $\mathbf{0}$ .

Les dérivées de Gâteaux sont en réalité surtout intéressantes lorsque des applications entre espaces de dimension infinie sont prises en compte. Il est alors même possible d'avoir une dérivée de Gâteaux linéaire sans que la dérivée de Fréchet existe...



# Chapitre 4

## Courbes

*“Il est important d’avoir de jolies courbes.”*

-LAURE ET AL

De manière assez naturelle, les courbes paramétrées apparaissent en tentant de modéliser le mouvement d’une particule (ou plus ambitieusement d’un corps). En effet, l’application qui à un temps  $t \in [t_1, t_2] \subset \mathbb{R}$  donne la position de la particule (*e.g.* dans l’espace tridimensionnel) n’est autre qu’une application de  $[t_1, t_2]$  dans  $\mathbb{R}^3$ . Les dimensions supérieures à 3 peuvent être atteintes lorsque plusieurs (disons  $N$ ) particules sont considérées (l’application associe alors  $N$  triplets de positions à un temps  $t$ , soit un espace d’arrivée de dimension  $3N$ ) ou encore dans le mouvement d’un corps rigide (où il faut ajouter à la position d’un point de base, quelques coordonnées d’angles).

Une courbe paramétrée n’est rien d’autre qu’une application  $\gamma(t) : I \rightarrow \mathbb{R}^n$ , où  $I \subset \mathbb{R}$  est *a priori* seulement un sous-ensemble de  $\mathbb{R}$ . Dans ce chapitre, les courbes paramétrées seront toujours supposées être des applications continues définies sur des intervalles. La trace ou l’image d’une courbe est l’ensemble  $\gamma(I) = \text{Im } \gamma \subset \mathbb{R}^n$ .

Lorsqu’une courbe paramétrée  $\gamma : I \rightarrow \mathbb{R}^n$  est dérivable en  $t_0 \in I$ , la notation

$$\gamma'(t_0) = \begin{pmatrix} \frac{dF_1}{dt}(t_0) \\ \frac{dF_2}{dt}(t_0) \\ \vdots \\ \frac{dF_n}{dt}(t_0) \end{pmatrix} = \gamma_*(t_0)$$

sera utilisée pour ce vecteur nommé vecteur vitesse de  $\gamma$  en  $t_0$ . Comme son nom l’indique, il représente la vitesse (vectorielle, *c’ad.* avec direction) instantanée d’une particule dont le mouvement dans  $\mathbb{R}^m$  serait décrit par  $\gamma$ . La vitesse de  $\gamma$  en  $t_0$  est la norme de  $\gamma'(t_0)$ , notée  $\|\gamma'(t_0)\|$ .

**Remarque 4.0.1.** Les propriétés obtenues précédemment pour les applications, restent valides ici. En particulier, soit  $I \subset \mathbb{R}$ ,  $\gamma, \sigma : I \rightarrow \mathbb{R}^n$ , toutes deux supposées dérivables, alors

1.  $(\gamma(t) + \sigma(t))' = \gamma'(t) + \sigma'(t)$  ;
2.  $(\gamma(t) \cdot \sigma(t))' = \gamma'(t) \cdot \sigma(t) + \gamma(t) \cdot \sigma'(t)$  ;
3.  $(\gamma(t) \wedge \sigma(t))' = \gamma'(t) \wedge \sigma(t) + \gamma(t) \wedge \sigma'(t)$  ;

4. pour une fonction  $f : S \rightarrow I$  (où  $S \subset \mathbb{R}$ ) dérivable  $(\gamma \circ f(t))' = \gamma'(f(t))f'(t)$  ;
5. pour une fonction dérivable  $f : I \rightarrow \mathbb{R}$ ,  $(f(t)\gamma(t))' = f'(t)\gamma(t) + f(t)\gamma'(t)$ .



Moult exemples de courbes (et plus de détails sur les courbes présentées ici) sont disponibles (parmi une foule de références possibles) sur le site <http://www.mathcurve.com/>

## 4.1 Paramétrage

Il est parfois aussi intéressant de regarder les courbes non pas en tant que mouvement mais pour la trace qu'elle décrit. En particulier, il peut être souhaitable de pouvoir donner un sens à la longueur d'une trace dans l'espace. Par exemple, le cercle est décrit par l'équation  $x^2 + y^2 = 1$ , c'est-à-d. c'est l'ensemble des points de  $\mathbb{R}^2$  satisfaisant cette relation. Mais il peut être paramétrée par une très grande famille de courbe, e.g.  $\mathbb{R} \ni t \mapsto (\sin t, \cos t)$ , ou  $\mathbb{R} \ni t \mapsto (\sin(t^2), \cos(t^2))$ , ou encore  $[-1, 1] \ni t \mapsto (1 - 2t, \mathbb{1}_{[0,1]}2\sqrt{t-t^2} - \mathbb{1}_{[-1,0]}2\sqrt{t-t^2})$ .

**Définition 4.1.1.** Le terme paramétrage d'une trace fait référence à une des courbes paramétrées qui peut décrire une même trace, i.e. pour un sous-ensemble  $S$  de  $\mathbb{R}^m$  fixé, un paramétrage est une application (la même ou une autre)  $\gamma : I \rightarrow \mathbb{R}^n$  telle que  $\text{Im } \gamma = S$  (si  $S$  est un ensemble tordu, il se peut qu'une telle application n'existe pas).

Le terme reparamétrage  $C^k$  (où  $k \in \mathbb{Z}_{\geq 0} \cup \{\infty, \omega\}$ ) fait référence à une courbe paramétrée  $\sigma : I_2 \rightarrow \mathbb{R}^m$  obtenue d'une courbe paramétrée  $\gamma : I_1 \rightarrow \mathbb{R}^m$  par  $\sigma = \gamma \circ f$  où  $f : I_2 \rightarrow I_1$  est bijective, de classe  $C^k$  et son inverse est de classe  $C^k$ . ★

La démonstration que la relation  $R_k(\gamma, \sigma)$  donnée par le fait que  $\sigma$  est un reparamétrage  $C^k$  de  $\gamma$  est une relation d'équivalence est laissée en exercice :

EXERCICE 55: Soit  $\gamma_{(1)} : I_1 \rightarrow \mathbb{R}^n$  et  $\gamma_{(2)} : I_2 \rightarrow \mathbb{R}^n$  deux courbes paramétrées. Soit  $k \in \mathbb{Z}_{\geq 0}$  fixé. Montrer que la relation définie par  $\gamma_{(1)} \simeq \gamma_{(2)}$  s'il existe un difféomorphisme de classe  $C^k$ ,  $f : I_2 \rightarrow I_1$  tel que  $\gamma_{(1)} \circ f = \gamma_{(2)}$  est une relation d'équivalence, i.e.

RE1.  $\gamma_{(1)} \simeq \gamma_{(1)}$  (réflexive)

RE2.  $\gamma_{(1)} \simeq \gamma_{(2)} \Rightarrow \gamma_{(2)} \simeq \gamma_{(1)}$  (symétrique)

RE3.  $\gamma_{(1)} \simeq \gamma_{(2)}$  et  $\gamma_{(2)} \simeq \gamma_{(3)} \Rightarrow \gamma_{(1)} \simeq \gamma_{(3)}$  (transitive)

En particulier, deux courbes paramétrées  $\gamma_{(1)} : I_1 \rightarrow \mathbb{R}^n$  et  $\gamma_{(2)} : I_2 \rightarrow \mathbb{R}^n$  sont souvent dites équivalentes s'il existe un reparamétrage  $C^1$ ,  $f : I_1 \rightarrow I_2$ , tel que  $\gamma_{(1)} \circ f = \gamma_{(2)}$ . Les exemples 4.1.4 et 4.1.7 montrent qu'il peut y avoir deux paramétrages de la même courbe qui ne sont pas des reparamétrages.

**Exemple 4.1.2.** Soit  $\varepsilon \in \mathbb{R}_{\geq 0}$ , soit  $m \in \mathbb{R}_{>0}$ , soit

$$I_\varepsilon = \{\theta \in [0, 2\pi[ \mid 1 - \varepsilon \cos \theta > 0 \text{ et } 1 + \varepsilon \cos \theta \geq 0\},$$

et soit

$$r(\theta) = \begin{cases} \frac{\varepsilon m}{1 - \varepsilon \cos \theta} & \text{si } \varepsilon \neq 0 \\ r(\theta) = m & \text{si } \varepsilon = 0. \end{cases}$$

Enfin, soit  $\gamma_{(\varepsilon)} : I_{\varepsilon} \rightarrow \mathbb{R}^2$  la courbe paramétrée définie par  $\gamma_{(\varepsilon)}(\theta) = \begin{pmatrix} r(\theta) \cos \theta \\ r(\theta) \sin \theta \end{pmatrix}$ . En fait, la trace des courbes  $\gamma_{(\varepsilon)}$  donne toute la famille des coniques (selon les valeurs du paramètre  $\varepsilon$ ). En particulier,  $\sigma : \mathbb{R} \rightarrow \mathbb{R}^2$  définie par  $\sigma(t) = \begin{pmatrix} t^2 - m^2 \\ t \end{pmatrix}$  est un reparamétrage  $C^1$  (i.e. un paramétrage équivalent) de la parabole  $\gamma_{(1)}$ . ♣

La forme du paramétrage vue dans l'exemple précédent est relativement fréquent, au point d'avoir un nom.

**Définition 4.1.3.** Un paramétrage en coordonnées polaires d'une trace est déterminé par une fonction  $r : I \rightarrow \mathbb{R}$  et défini par  $\gamma(\theta) = (r(\theta) \cos \theta, r(\theta) \sin \theta)$ . Il est commun de demander de plus que  $\forall \theta \in I, r(\theta) \geq 0$ . ★

Outre de donner un exemple de paramétrage polaire, le prochain exemple reviendra plus tard pour montrer que des courbes paramétrées qui ont la même trace peuvent être néanmoins différentes.

**Exemple 4.1.4.** Soit  $r(t) = \sqrt{|\cos(2t)|}$  définie sur  $[0, 2\pi]$ , alors la courbe paramétrée en coordonnées polaires

$$\gamma(t) = \begin{pmatrix} \sqrt{|\cos(2t)|} \cos t \\ \sqrt{|\cos(2t)|} \sin t \end{pmatrix}$$

dessine un échangeur d'autoroute (l'échangeur d'autoroute possède cette forme car la courbe a des propriétés particulières). Cette courbe est appelée la lemniscate de Bernoulli, c'est aussi les points  $(x, y) \in \mathbb{R}^2$  solutions de  $(x^2 + y^2)^2 = \pm(x^2 - y^2)$ . [Le domaine de définition est parfois restreint aux angles où  $\cos 2t > 0$ , dans quel cas il s'agit des points solution de  $(x^2 + y^2)^2 = (x^2 - y^2)$ .] ♣

Lorsqu'une courbe paramétrée  $\gamma : I \rightarrow \mathbb{R}^n$  est telle que  $\gamma'(t) = \mathbf{0}$  pour un  $t \in I$ , il est possible (un peu comme dans le cas du gradient) que ce soit le signe d'un comportement étrange de la trace.

**Exemple 4.1.5.** Soit  $\gamma(t) = \begin{pmatrix} t^3/3 \\ t^2/2 \end{pmatrix}$  une courbe paramétrée de classe  $C^\infty$ . Alors,  $\gamma'(t) = \begin{pmatrix} t^2 \\ t \end{pmatrix}$  s'annule en  $t = 0$ . En fait, tout paramétrage de classe  $C^1$  de cette trace aura la même propriété...

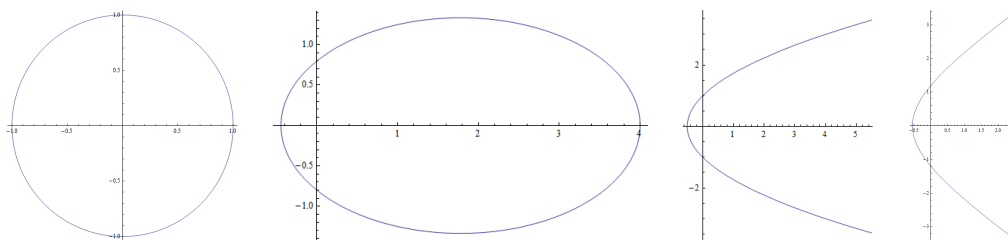


FIGURE 4.1 – Voici (une partie de) la trace de quelques unes de ces courbes paramétrées de l'exemple 4.1.2 (pour  $m = 1$  et, de gauche à droite,  $c = 0, \frac{4}{5}, 1$  et  $\frac{6}{5}$ ).

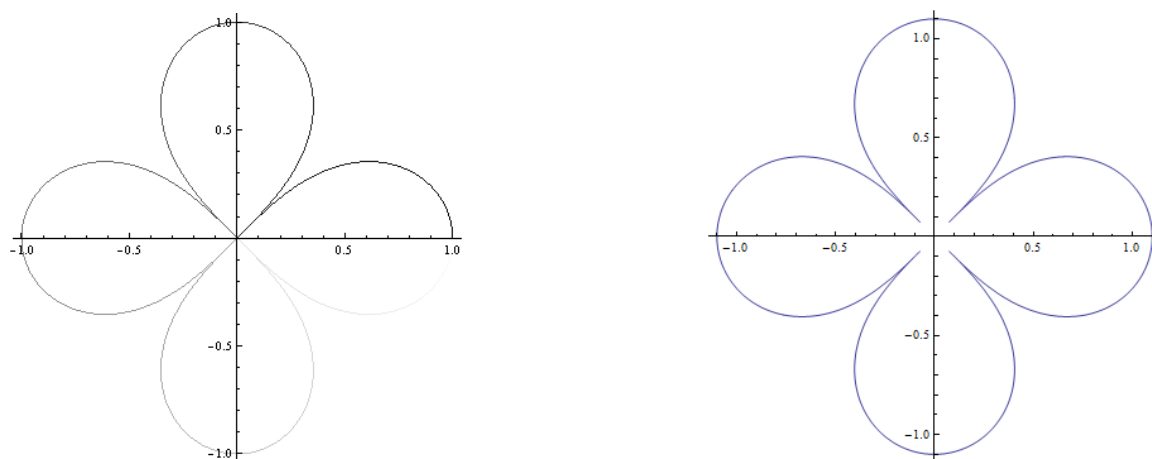


FIGURE 4.2 – Le paramétrage polaire de la lemniscate (exemple 4.1.4) : à gauche le trait est noir en  $t = 0$  puis s'éclaircit lorsque  $t$  grandit, à droite elle est un peu déformée pour mettre en exergue qu'il y a des rebroussements.

Pour le voir, il suffit de remarquer qu'un paramétrage  $C^0$  de cette trace passera tôt ou tard de la partie gauche (où la première coordonnée est négative) à la partie droite (celle où elle est positive), en passant par le point  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  (la projection sur la première coordonnée est une bijection de la trace de  $\gamma$  sur  $\mathbb{R}$  ; il suffit de voir que  $\pi_1 \circ \gamma(t) = t^3/3$  est une bijection). De plus, le premier paramétrage  $\gamma$  permet quand même de dire que le plan tangent à la courbe en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est l'axe  $x = 0$ , en prenant la limite des plans tangents pour  $t \neq 0$ , *i.e.* comme tout vecteur vitesse d'un paramétrage est dans le plan tangent (*cf.* théorème 3.7.9), si le paramétrage est de classe  $C^1$  la limite des vecteurs vitesses appartiendra à la limite des vecteurs tangents.

Soit  $t_0$  tel que  $\phi(t_0) = \mathbf{0}$ . Soit  $I'$  un voisinage de  $t_0$  tel que  $\phi(I') = B_\varepsilon(\mathbf{0}) \cap \text{Im } \phi$ , *i.e.*  $\phi$  restreinte à  $I'$  est bijective sur un voisinage de la trace en l'origine. Lorsque  $t$  varie en croissant sur  $I'$ ,  $\phi(t)$  va soit de gauche à droite, soit de droite à gauche. Dans les deux cas, la limite du vecteur vitesse pour  $t < t_0$  et pour  $t > t_0$  existe (comme le paramétrage est supposé  $C^1$ ), mais ces deux limites sont forcément de signes opposés (dans un cas la seconde coordonnée croît tandis que dans l'autre elle décroît). Il faut donc que le vecteur vitesse s'annule en ce point.

Avant de clore cet exemple, il est bon de mentionner le paramétrage  $\sigma(t) = \begin{pmatrix} t/3 \\ |t|^{2/3}/2 \end{pmatrix}$  qui est de classe  $C^0$  n'est jamais de vitesse nulle, puisque la dérivée n'est pas définie en  $t = 0$  (ce qui ne contredit rien de ce qui est écrit plus haut puisque la vitesse n'est pas non plus définie en ce point). ♣

Pour parler de ce phénomène pour une trace, il est non seulement nécessaire de regarder tous les paramétrages, mais il est parfois nécessaire, pour recouvrir un petit voisinage d'un point  $\mathbf{x}$  d'une trace, d'avoir à faire à plusieurs courbes paramétrées (*cf.* dans l'exemple 4.1.4 prendre un petit voisinage de l'origine coupe forcément la courbe paramétrée originale).

**Définition 4.1.6.** Une courbe paramétrée  $\gamma : I \rightarrow \mathbb{R}^n$  est dite régulière si elle est de classe  $C^1$  et  $\forall t \in \mathring{I}, \gamma'(t) \neq \mathbf{0}$  (où  $\mathring{I}$  est l'intérieur de  $I$ ).

Un point  $\mathbf{x} \in \text{Im} \gamma$  est un point singulier de cette courbe paramétrée si le vecteur vitesse en ce point est nul pour tout reparamétrage. ★

Finalement, le prochain exemple montre qu'il y a une différence non triviale entre reparamétrage et paramétrage (et que les points singuliers peuvent être plus subtils qu'ils n'en ont l'air).

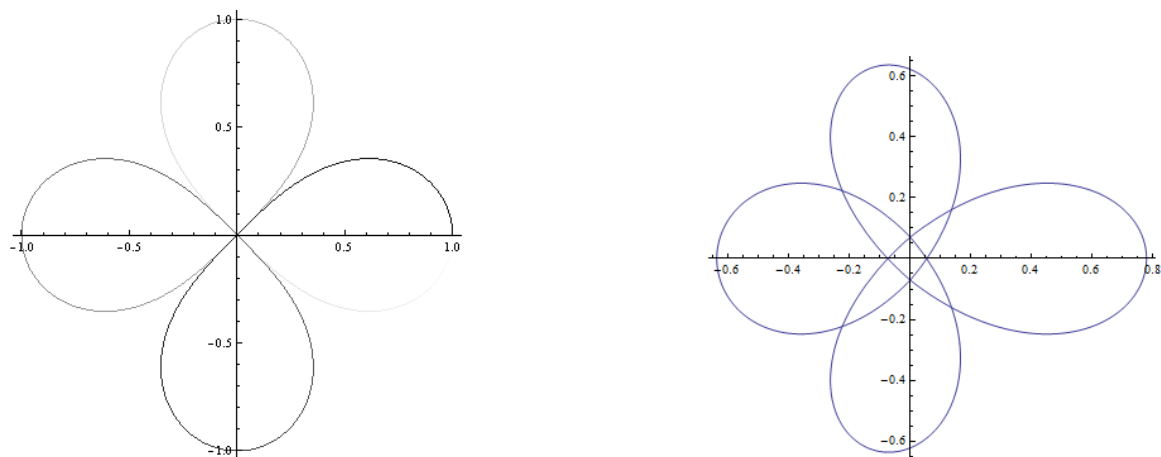


FIGURE 4.3 – Le paramétrage régulier de la lemniscate (exemple 4.1.7) : à gauche en ton de plus en plus clairs quand le temps augmente, à droite déformé.

**Exemple 4.1.7.** La courbe paramétrée  $\sigma : [0, 2\pi] \rightarrow \mathbb{R}^2$  définie par

$$\sigma(t) = \begin{pmatrix} \frac{\cos(2t)}{\sqrt{|\cos(2t)|}} \cos t \\ \frac{\cos(2t)}{\sqrt{|\cos(2t)|}} \sin t \end{pmatrix}$$

aura aussi pour trace la lemniscate de l'exemple 4.1.4. Avec ce paramétrage la vitesse est parfois  $\infty$ , mais il est possible de trouver un autre paramétrage qui est une courbe paramétrée régulière. Ainsi, lorsqu'il y a des points d'auto-intersections, deux courbes paramétrées de même trace ne sont pas forcément un reparamétrage l'une de l'autre. ♣

**Définition 4.1.8.** Une division d'un intervalle est la donnée d'une suite finie  $t_i \in [a, b]$  telle que  $a = t_1 < t_2 < \dots < t_n = b$ . La largeur d'une division est alors  $\delta(\{t_i\}) = \max_{i \in \bar{n}} (t_{i+1} - t_i)$ . ★

**Définition 4.1.9.** Une courbe paramétrée  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  est dite  $C^1$  par morceaux (resp. régulière par morceaux) si elle est  $C^0$  sur  $I$  et qu'il existe une division  $\{t_i\}_{i \in \bar{n}}$  telle que

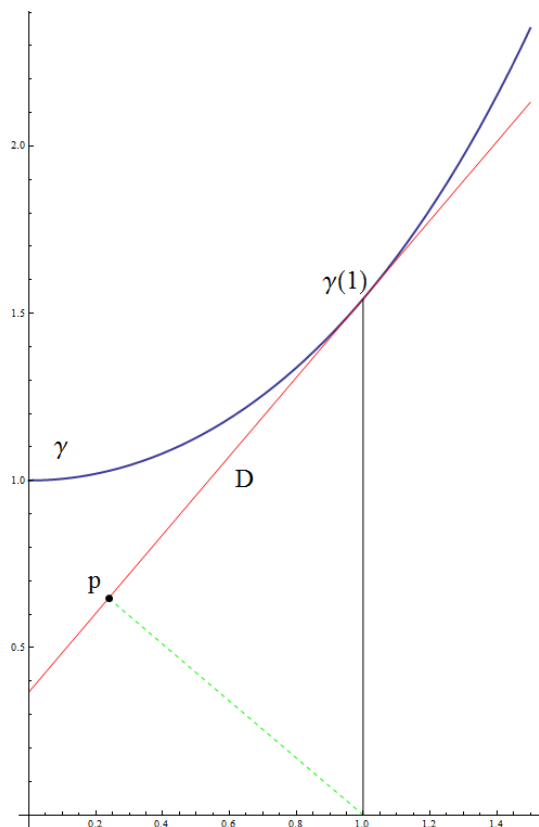
- $\gamma$  est  $C^1$  (resp. régulière) sur chaque intervalle ouvert  $]t_i, t_{i+1}[$
- et que  $\forall i \in \bar{n}$  les deux limites  $\lim_{t \rightarrow t_i^\pm} \gamma'(t)$  existent (mais ne sont pas nécessairement égales !).

★

Ainsi, une courbe  $C^1$  par morceaux peut-être vue comme plusieurs courbes  $C^1$  mises bout à bout. Pour avoir vraiment toutes les propriétés des courbes, il est normalement nécessaire de parler en terme de courbes paramétrées régulières. Cependant les courbes paramétrées  $C^1$  par morceaux apparaissent naturellement, e.g. le bord d'un pavé de  $\mathbb{R}^2$  est la trace d'une courbe paramétrée  $C^1$  par morceaux.

## 4.2 - Longueur ; cas Lipschitz

EXERCICE 56: Étant donné une (demi-)chaînette (i.e. la courbe paramétrée  $\gamma: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = (t, \cosh t)$ ), le but de cet exercice est de trouver une “développante”. La construction va comme suit. Soit donné un point  $\gamma(t) \in \text{Im } \gamma$ .



- Donner l'équation de la droite  $D$  tangente en ce point (s'assurer que la droite passe bien par  $\gamma(t)$ ).
- Trouver le point  $\mathbf{p}$  de  $D$  le plus proche de  $(\gamma_1(t), 0) = (t, 0)$  et évaluer la distance de  $\mathbf{p}$  à  $(t, 0)$ . [Indice : trouver un vecteur  $\mathbf{v}$  orthogonal au vecteur directeur de  $D$ , puis chercher  $\lambda$  tel que  $\lambda \mathbf{v} + (t, 0) \in D$ .]
- Écrire l'expression de  $\mathbf{p}$  en fonction de  $t$  (c'ad. en fonction de  $\gamma(t)$ ).
- La courbe paramétrée obtenue est la tractrice. C'est la courbe décrite par un objet  $a$  placé en  $(0, 1)$  qui est relié par une tige (rigide mais qui pivote aux deux extrémités) de longueur 1 à un objet  $b$  (placé en  $(0, 0)$ ) lorsque l'objet  $b$  se déplace le long de l'axe des  $x$  vers la droite. Qu'est ce qui dans la construction ci-dessus vous permet de le dire ?

EXERCICE 57: Dans un fameux épisode de Sherlock Holmes, notre détective favori tente de décider, en regardant les traces qu'un vélo laisse sur le sol, dans quelle direction ce vélo allait. Quatre hypothèses sont faites :

- 1- les deux roues restent toujours au sol,    2- le vélo doit toujours être en mouvement,
- 3- la roue avant est la roue directrice, et    4- la roue arrière est la roue tractrice.

Quelles courbes peuvent être tracées par la roue avant mais pas par la roue arrière ? Pourriez-vous exprimer l'une en fonction de l'autre ? Que se passe-t-il si les hypothèses sont modifiées ?

## 4.2 Longueur ; cas Lipschitz

Une des quantités importantes qu'il est souhaitable d'associer à une courbe est sa longueur<sup>1</sup>. Le principe de base ici est le théorème de Pythagore (ou plutôt la notion de distance) : soit  $a, b \in \mathbb{R}_{>0}$ , et soit  $\gamma: [0, 1] \rightarrow \mathbb{R}^2$  la courbe paramétrée définie par  $\gamma(t) = (at, bt)$ , il faudrait que la longueur de  $\gamma$  soit  $\sqrt{a^2 + b^2}$ .

1. Tout ce qui se trouve après la proposition 4.2.3 est pour la curiosité du lecteur uniquement.



Pour parler de la longueur d'une courbe paramétrée  $\gamma : [a, b] \rightarrow \mathbb{R}^m$ , la courbe sera approximée par des segments (alias une courbe polygonale). Si l'approximation est bonne (pour une certaine famille), la longueur pourra être définie.

**Définition 4.2.1.** Soit  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  une courbe paramétrée ( $C^0$ ), sa longueur [d'arc] est

$$\mathcal{L}ong(\gamma) = \sup_{a=t_1 < \dots < t_n=b} \sum_{i=1}^{n-1} \|\gamma(t_i) - \gamma(t_{i+1})\|,$$

le supremum étant pris sur toutes les divisions de l'intervalle  $[a, b]$  (en  $n - 1$  morceaux,  $n$  pouvant être aussi grand que voulu).

La courbe paramétrée  $\gamma$  est dite rectifiable si le supremum ci-dessus (sa longueur) est fini. ★

La première remarque à faire à partir de cette définition, est qu'elle invariante sous reparamétrage  $C^0$ . À partir de là, il n'est pas très dur (exercice) de montrer que, si la courbe n'a qu'un nombre fini d'intersection avec elle-même, elle ne dépend pas du paramétrage.

Notons que  $\gamma$  étant continue et définie sur un compact, son image est bornée. Une suite de divisions  $\{t_{i,(k)}\}_{i \in \overline{n_k}}$  de l'intervalle qui donne une approximation par des segments de longueur totale de plus en plus en grande est forcément une suite de divisions qui contient des bouts de plus en plus petits, i.e.  $\lim_{k \rightarrow \infty} \min_{i \in \overline{n_k}} |t_{i+1,(k)} - t_{i,(k)}| = 0$ .

Pour une suite croissante  $\{t_i\}_{i \in \overline{n}}$  d'éléments de  $[a, b]$  (i.e.  $t_1$  n'est pas nécessairement  $a$  et  $t_n$  n'est pas nécessairement  $b$ ), introduisons la notation

$$\Sigma_\gamma(\{t_i\}) = \sum_{i=1}^{n-1} \|\gamma(t_i) - \gamma(t_{i+1})\|.$$

**Remarque 4.2.2.** Grâce à l'inégalité du triangle, il est facile (mais très important) de constater que si une division  $\{t'_i\}$  est obtenue d'une division précédente  $\{t_i\}$  en coupant un peu plus finement (i.e. en ajoutant un ou plusieurs réels dans la suite  $\{t_i\}$ ) alors  $\Sigma_\gamma(\{t_i\}) \leq \Sigma_\gamma(\{t'_i\})$ . Autrement dit, couper plus finement rapproche toujours du supremum dans la définition 4.2.1 de la longueur. Ceci ne veut pas pour autant dire que n'importe quelle division avec beaucoup de points est bonne. ♠

Rappelons qu'une application  $\mathbf{F} : S \rightarrow \mathbb{R}^m$  est Lipschitz si  $\exists K \in \mathbb{R}_{>0}$  tel que  $\forall \mathbf{x}, \mathbf{y} \in S, \|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\| \leq K \|\mathbf{x} - \mathbf{y}\|$ . Lorsque la constante  $K$  est connue, le terme  $K$ -Lipschitz est alors employé.

**Proposition 4.2.3**

Si  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  est Lipschitz, alors  $\gamma$  est rectifiable.

DÉMONSTRATION: Si  $\gamma$  est Lipschitz et soit  $K \in \mathbb{R}_{>0}$  alors, pour toute division de  $[a, b]$

$$\Sigma_\gamma(\{t_i\}) = \sum_{i=0}^{n-1} \|\gamma(t_i) - \gamma(t_{i+1})\| \leq \sum_{i=0}^{n-1} K |t_i - t_{i+1}| \leq K(b - a).$$

Ainsi le supremum est forcément fini, puisque c'est le supremum d'un ensemble contenu dans  $[0, K(b - a)]$ . ■

Il est relativement facile de construire une fonction qui est Lipschitz mais non dérivable sur un ensemble dense (e.g. les fractions binaires). À ce titre, il est justifié de se demander si ce critère est complet, c'est-à-dire. une courbe paramétrée rectifiable est-elle Lipschitz. Le théorème suivant est présenté pour la curiosité du lecteur :

**Théorème 4.2.4**

*Si  $\gamma$  est rectifiable, alors il existe un reparamétrage  $C^0$  de  $\gamma$  qui est un paramétrage Lipschitz de  $\text{Im } \gamma$ , i.e.  $\exists f : [0, L'] \rightarrow [a, b]$  un difféomorphisme  $C^0$  (alias un homéomorphisme, alias une application bicontinue) tel que  $\gamma \circ f$  est une courbe Lipschitz (de même trace que  $\gamma$ ).*

Ainsi, une courbe est rectifiable si et seulement si il existe un reparamétrage  $C^0$  qui est Lipschitz.

DÉMONSTRATION: Supposons que  $\gamma$  est rectifiable et soit  $L = \text{Long}(\gamma)$ . Alors  $\forall c \in [a, b]$ ,  $\gamma|_{[a, c]}$  est aussi rectifiable. Soit  $\mathcal{A}(c) = \text{Long}(\gamma|_{[a, c]})$  la longueur de ce sous-arc de la courbe.

La démonstration de la continuité de  $\mathcal{A}$  sera d'abord faite. Supposons que  $c_k$  est une suite telle que  $\lim_{k \rightarrow \infty} c_k = c$  et  $|\lim_{k \rightarrow \infty} \mathcal{A}(c_k) - \mathcal{A}(c)| > \varepsilon$ . Par continuité de  $\gamma$ , soit  $\delta$  tel que  $\forall t$  tel que  $|t - c| < \delta$ ,  $\|\gamma(t) - \gamma(c)\| < \varepsilon/60$ . Puis par définition de la limite, soit  $k_0$  tel que  $\forall k \geq k_0$ ,  $|c_k - c| < \delta$ . Enfin, soit  $c_M = \max(\sup_{k \geq k_0} c_k, c)$  et  $c_m = \min(\inf_{k \geq k_0} c_k, c)$ .

Montrons d'abord qu'une bonne division de  $[a, c_M]$  donne lieu à une bonne division de tous les  $[a, c_k]$  ( $k \geq k_0$ ) et de  $[a, c]$ . Soit  $\{t_i\}_{i \in \overline{n}}$  une division telle que  $\mathcal{A}(c_M) - \Sigma_\gamma(\{t_i\}) < \frac{\varepsilon}{6}$  et que cette division contienne  $c_m$  et  $c$ . Soit  $\{t'_{i,k}\}_{i \in \overline{n'}} = \{t_i\}_{i \in \overline{n}} \cap [a, c_k] \cup \{c_k\}$  (la même construction se fait pour  $c$ ). [Par contradiction] Supposons que  $\mathcal{A}(c_k) - \Sigma_\gamma(\{t'_{i,k}\}) > \frac{\varepsilon}{4}$ , alors il doit exister en particulier une division  $\{t''_i\}$  de  $[a, c_k]$  telle que  $\Sigma_\gamma(\{t''_i\}) - \Sigma_\gamma(\{t'_{i,k}\}) > \frac{\varepsilon}{4}$  et que cette division contient  $c$ . Alors, comme en général  $\forall s' \in \{s_i\}$  un découpage de  $[a, c_M]$ ,  $\Sigma_\gamma(\{s_i\}) = \Sigma_\gamma(\{s_i\} \cap [a, s']) + \Sigma_\gamma(\{s_i\} \cap [s', c_M])$ , il apparaît que si  $t_{M,k}$  est le plus grand des  $t_i$  tel que  $t_i < c_k$ ,

$$\begin{aligned} |\Sigma_\gamma(\{t''_i\} \cup \{t_i\}) - \Sigma_\gamma(\{t_i\})| &= |\Sigma_\gamma(\{t''_i\} \cup \{t_i\}) - \Sigma_\gamma(\{t'_{i,k}\}) + \Sigma_\gamma(\{t'_{i,k}\}) - \Sigma_\gamma(\{t_i\})| \\ &\geq |\Sigma_\gamma(\{t''_i\} \cup \{t_i\}) - \Sigma_\gamma(\{t'_{i,k}\}) - \Sigma_\gamma(\{t_i\} \cap [t_{M,k}, c_M])| \\ &\quad - |\Sigma_\gamma(\{t'_{i,k}\}) - \Sigma_\gamma(\{t_i\} \cap [a, t_{M,k}])| \\ &\geq |\Sigma_\gamma(\{t''_i\} \cup \{t_i\}) - \Sigma_\gamma(\{t_i\}) - \|\gamma(c_k) - \gamma(t_{M,k})\|| - \|\gamma(c_k) - \gamma(t_{M,k})\| \\ &\geq |\Sigma_\gamma(\{t''_i\} \cup \{t_i\}) - \Sigma_\gamma(\{t_i\})| - \|\gamma(c_k) - \gamma(t_{M,k})\| \\ &\quad - \|\gamma(c_k) - \gamma(c) + \gamma(c) - \gamma(t_{M,k})\| \\ &> \frac{\varepsilon}{4} - 4 \frac{\varepsilon}{60} = \frac{11\varepsilon}{60} \end{aligned}$$

car  $\Sigma_\gamma(\{t_i\})$  et  $\Sigma_\gamma(\{t'_{i,k}\})$  ne diffèrent que par la présence de  $c_k$  (et comme  $c_k \in [c_m, c_M]$  et que  $c_m$  appartient à la division  $\{t_i\}$ ,  $\gamma(c_k)$  et  $\gamma(t_{M,k})$  sont  $\frac{\varepsilon}{60}$  proches de  $\gamma(c)$ ). Pour résumer, si les approximations  $\Sigma_\gamma(\{t'_{i,k}\})$  sont trop mauvaises ( $> \frac{\varepsilon}{4}$ ), l'approximation  $\Sigma_\gamma(\{t_i\})$  aurait un écart de  $> \frac{11\varepsilon}{60}$  avec  $\Sigma_\gamma(\{t''_i\} \cup \{t_i\})$  qui est à son tour à un certain écart (peut-être nul) du supremum  $\mathcal{A}(c_M)$ . Or  $\Sigma_\gamma(\{t_i\})$  a été choisie parce qu'elle présentait un écart  $< \frac{10}{60}\varepsilon$  de  $\mathcal{A}(c_M)$ , d'où la contradiction. Ainsi  $\mathcal{A}(c_k) - \Sigma_\gamma(\{t'_{i,k}\}) \leq \frac{\varepsilon}{4}$ .

Maintenant, comme  $c_k \rightarrow c$ , lorsque  $k$  est assez grand,  $\Sigma_\gamma(\{t'_{i,k}\}) = \Sigma_\gamma(\{t_i\} \cap [a, c]) \pm \|\gamma(c_k) -$

$\gamma(c)\|$ , car il n'y a plus d'éléments de  $\{t_i\}$  entre  $c$  et  $c_k$  (le signe dépend du signe de  $c_k - c$ ). Ainsi,

$$\begin{aligned} |\mathcal{A}(c_k) - \mathcal{A}(c)| &= |\mathcal{A}(c_k) - \Sigma_\gamma(\{t'_i\}) + \Sigma_\gamma(\{t'_i\}) - \Sigma_\gamma(\{t_i\} \cap [a, c]) + \Sigma_\gamma(\{t_i\} \cap [a, c]) - \mathcal{A}(c)| \\ &\leq |\mathcal{A}(c_k) - \Sigma_\gamma(\{t'_i\})| + |\Sigma_\gamma(\{t'_i\}) - \Sigma_\gamma(\{t_i\} \cap [a, c])| + |\Sigma_\gamma(\{t_i\} \cap [a, c]) - \mathcal{A}(c)| \\ &\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{60} + \frac{\varepsilon}{4} = \frac{31}{60}\varepsilon < \varepsilon \end{aligned}$$

D'où la contradiction.  $\mathcal{A} : [a, b] \rightarrow [0, L]$  est donc continue.

$\mathcal{A} : [a, b] \rightarrow [0, L]$  est (non-strictement) croissante, ainsi, elle peut-être constante sur  $[s, t] \subset [a, b]$  (cela correspond à un temps où le paramétrage s'arrête). La préimage de  $\mathbf{x} \in \text{Im } \gamma$  est une union disjointe de points et d'intervalles. Ainsi, les endroits où  $\mathcal{A} : [a, b] \rightarrow [0, L]$  ne croît pas (*i.e.* reste constante) sont des intervalles disjoints, disons  $\cup [a_i, b_i]$  (il peut y en avoir une infinité dénombrable, mais la série  $\sum_i (b_i - a_i)$  est forcément convergente). Il faut alors modifier un peu  $\mathcal{A}$  sur ces intervalles pour qu'elle devienne strictement croissante. Soit

$$\mathcal{A}'(c) = \mathcal{A}(c) + \max_{b_i > t} (0, t - a_i) + \sum_{b_i \leq t} (b_i - a_i).$$

Alors  $\mathcal{A}' : [a, b] \rightarrow [0, L]$  est strictement croissante (puisqu'elle vient d'une fonction croissante qui a été modifiée pour ne plus être constante). Elle est de plus continue, donc sa réciproque  $f = \mathcal{A}'^{-1}$  est aussi continue (et strictement croissante). Il reste à montrer que  $\gamma \circ f$  est Lipschitz. Or, l'arc de  $\gamma \circ f(s)$  à  $\gamma \circ f(s')$  (disons que  $s < s'$ ) se caractérise par le fait que toute subdivision est la longueur inférieure ou égale à  $s' - s$  (strictement inférieure s'il s'agit d'une région où la courbe s'arrête). En particulier, le segment direct entre les deux points (la division "triviale" qui est faite en ne regardant que les deux extrémités) le sera. Ainsi,  $\|\gamma \circ f(s) - \gamma \circ f(s')\| \leq \|s - s'\|$ . Autrement dit, ce paramétrage est 1-Lipschitz. ■

En fait, comme  $\mathcal{A}$  est croissante, le seul souci qui pourrait arriver dans son inversion est un "plateau" dans son graphe, *i.e.* un endroit où elle ne croît pas. À toute fin pratique, il est possible de prendre  $f$  directement comme une réciproque de  $\mathcal{A}$  (*c'ad.* une fonction dont le graphe la réflexion par rapport à la droite  $x = y$  du graphe de  $\mathcal{A}$ , et, aux endroits où un plateau a été transformé en un segment vertical,  $f$  prend comme valeur n'importe quel point de ce segment) alors  $\gamma \circ f$  est exactement 1-Lipschitz partout, *i.e.*  $\forall t \in [0, L], v(t) := \limsup_{h \rightarrow 0} \frac{\|\gamma \circ f(t+h) - \gamma(t)\|}{h} = 1$ .

Ceci motive de regarder des courbes qui ne font pas de "pauses" ou qui ne s'arrêtent pas en un point pendant un certain temps.

**Définition 4.2.5.** Une courbe  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  est dite monotone si  $\forall c \in [a, b], \exists \varepsilon \in \mathbb{R}_{>0}$  tel que  $\gamma|_{]c-\varepsilon, c+\varepsilon[}$  est injective. ★

En refaisant la démonstration du théorème 4.2.4 pour une courbe monotone, il apparaîtra que la fonction  $\mathcal{A}$  est strictement croissante. Le reparamétrage est alors directement donné par l'inverse de cette fonction (il sera rencontré de nouveau sous des atours plus accorts).

### 4.3 Longueur ; cas $\mathcal{C}^1$ par morceaux

Une courbe paramétrée qui est  $\mathcal{C}^1$  par morceaux, aura sa dérivée bornée. Par le théorème des accroissements finis 6.1.3, elle est ainsi Lipschitz, et en particulier de longueur finie. Il est possible

### 4.3 - Longueur ; cas $C^1$ par morceaux

de relier plus fortement la dérivée à la longueur. Une petite variation de  $t$  donne lieu à une variation  $\gamma'(t)$  de la courbe. Or la longueur de  $\gamma'(t)$  est  $\|\gamma'(t)\|$ . Pour avoir la longueur totale, l'intuition indique qu'il suffit de faire la somme de ces petites variations, autrement dit l'intégrale de  $\|\gamma'(t)\|$ . Un lemme préparatoire est nécessaire :

#### Lemme 4.3.1

Soit  $I$  un intervalle et  $\gamma : I \rightarrow \mathbb{R}^n$  une courbe paramétrée alors

$$\left\| \int_I \gamma' \right\| \leq \int_I \|\gamma'\|,$$

où l'intégrale de gauche est une intégrale coordonnée par coordonnée (c'ad. elle a pour résultat le vecteur dont la  $i^{\text{ème}}$  coordonnée est  $\int_I \gamma'_i$ ).

La démonstration est remise au chapitre 7, voir lemme 7.3.8

#### Proposition 4.3.2

Soit  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  une courbe paramétrée  $C^1$  par morceaux, alors

$$\mathcal{L}ong(\gamma) = \int_a^b \|\gamma'(t)\| dt,$$

DÉMONSTRATION: Si la courbe n'est que  $C^1$  par morceaux, il suffit de regarder ce qui se passe sur chaque morceau ; en effet, il est toujours possible de supposer que la division qui est employée pour s'approcher du supremum, contient les points où la courbe n'est pas  $C^1$ . Par conséquent, il sera supposé que  $\gamma$  est de classe  $C^1$ .

Il sera démontré que  $\forall \varepsilon > 0, \left| \mathcal{L}ong(\gamma) - \int_I \|\gamma'(t)\| dt \right| < \varepsilon$ . Soit fixé un  $\varepsilon$ .

Tout d'abord, comme  $\gamma'$  est continue sur  $[a, b]$  (un compact), elle est uniformément continue. Ainsi,  $\forall \varepsilon, \exists \delta'$  tel que  $|t - s| < \delta' \Rightarrow \|\gamma'(s) - \gamma'(t)\| < \frac{\varepsilon}{4(b-a)}$ .

En particulier,  $\|\gamma'\|$  sera aussi uniformément continue, avec le même paramètre :  $|t - s| < \delta' \Rightarrow \left| \|\gamma'(s)\| - \|\gamma'(t)\| \right| < \frac{\varepsilon}{4(b-a)}$ . Par définition de l'intégrale de Riemann (en une variable) et la preuve du lemme 4.3.1, si  $\{t_i\}_{i \in \overline{n}}$  est une division de  $[a, b]$  telle que  $\delta(\{t_i\}) < \delta'$ , alors

$$\left| \int_a^b \|\gamma'(t)\| dt - \sum_{i=1}^{n-1} (t_{i+1} - t_i) \sup_{t \in [t_i, t_{i+1}]} \|\gamma'(t)\| \right| < \varepsilon/4$$

Ensuite, soit  $\{t_i\}_{i \in \overline{n}}$  une division de  $[a, b]$  telle que sa largeur  $\delta(\{t_i\})$  est  $< \delta'$  et que  $\mathcal{L}ong(\gamma) - \Sigma_{\gamma}(\{t_i\}) < \varepsilon/4$  (il suffit de prendre une division qui donne une bonne approximation [qui existe par

définition du sup] puis de lui rajouter des points [ce qui ne peut que la rapprocher du sup]). Alors

$$\begin{aligned}
 |\mathcal{L}ong(\gamma) - \int_a^b \|\gamma'(t)\| dt| &\leq |\mathcal{L}ong(\gamma) - \Sigma_\gamma(\{t_i\})| + \left| \sum_{i=1}^{n-1} \|\gamma(t_i) - \gamma(t_{i+1})\| - \sum_{i=1}^{n-1} \|\gamma'(t_i)\| (t_{i+1} - t_i) \right| \\
 &\quad + \left| \sum_{i=1}^{n-1} \|\gamma'(t_i)\| (t_{i+1} - t_i) - \sum_{i=1}^{n-1} (t_{i+1} - t_i) \sup_{t \in [t_i, t_{i+1}]} \|\gamma'(t)\| \right| \\
 &\quad + \left| \sum_{i=1}^{n-1} (t_{i+1} - t_i) \sup_{t \in [t_i, t_{i+1}]} \|\gamma'(t)\| - \int_a^b \|\gamma'(t)\| dt \right| \\
 &< \frac{\varepsilon}{4} + \sum_{i=1}^{n-1} \left| \|\gamma(t_i) - \gamma(t_{i+1})\| - \|\gamma'(t_i)\| (t_{i+1} - t_i) \right| \\
 &\quad + \sum_{i=1}^{n-1} (t_{i+1} - t_i) \left| \|\gamma'(t_i)\| - \sup_{t \in [t_i, t_{i+1}]} \|\gamma'(t)\| \right| + \frac{\varepsilon}{4} \\
 &< \frac{2\varepsilon}{4} + \sum_{i=1}^{n-1} \left| \|\gamma(t_i) - \gamma(t_{i+1})\| - \|(t_{i+1} - t_i)\gamma'(t_i)\| \right| + \sum_{i=1}^{n-1} (t_{i+1} - t_i) \frac{\varepsilon}{4(b-a)} \\
 &< \frac{3\varepsilon}{4} + \sum_{i=1}^{n-1} \left| \|\gamma(t_i) - \gamma(t_{i+1})\| - \|(t_{i+1} - t_i)\gamma'(t_i)\| \right|
 \end{aligned}$$

Pour le dernier terme, il faut faire un peu attention. Tout d'abord  $\gamma(t_{i+1}) - \gamma(t_i) = \int_{t_i}^{t_{i+1}} \gamma'(s) ds$  (où l'intégrale est à prendre, comme précédemment, coordonnée par coordonnée). Par conséquent

$$\gamma(t_{i+1}) - \gamma(t_i) = (t_{i+1} - t_i)\gamma'(t_i) + \int_{t_i}^{t_{i+1}} (\gamma'(s) - \gamma'(t_i)) ds.$$

Si  $\mathbf{b}_i = \int_{t_i}^{t_{i+1}} (\gamma'(s) - \gamma'(t_i)) ds$ , alors  $\|\mathbf{b}_i\| < (t_{i+1} - t_i) \frac{\varepsilon}{4(b-a)}$  (par continuité uniforme). En notant,  $\mathbf{a}_i = \|(t_{i+1} - t_i)\gamma'(t_i)\|$ , il apparaît que  $\|\gamma(t_i) - \gamma(t_{i+1})\| = \|\mathbf{a}_i + \mathbf{b}_i\|$ . Le dernier terme à majorer est une somme de  $\|\mathbf{a}_i + \mathbf{b}_i\| - \|\mathbf{a}_i\|$ , dont chacun est  $\leq \|\mathbf{b}_i\|$ . D'où

$$|\mathcal{L}ong(\gamma) - \int_a^b \|\gamma'(t)\| dt| < \frac{3\varepsilon}{4} + \sum_{i=1}^{n-1} (t_{i+1} - t_i) \frac{\varepsilon}{4(b-a)} = \varepsilon. \quad \blacksquare$$

**Remarque 4.3.3.** La démonstration ci-dessus peut s'étendre au cas où  $\gamma$  n'est que Lipschitz. En effet, il faut alors définir  $v_\gamma(t) = \limsup_{h \rightarrow 0} \frac{\|\gamma(t+h) - \gamma(t)\|}{h}$ . Le résultat obtenu est alors que  $\mathcal{L}ong(\gamma) = \int_a^b v(t) dt$ . La définition faite ici permet en fait de parler de courbe dans un espace métrique général. ♠

**Remarque 4.3.4.** Il est malheureusement commun de voir des “démonstrations” du théorème 4.3.2 faire intervenir le théorème 3.1.6 des valeurs intermédiaires. Ce résultat (une conséquence du lemme de Rolle) sur les fonctions de  $\mathbb{R} \rightarrow \mathbb{R}$  dit que si  $f$  est une fonction continue sur l'intervalle  $[a, b]$  et dérivable sur l'intérieur de cet intervalle, alors  $\exists s \in [a, b]$  tel que  $f(b) - f(a) = (b - a)f'(s)$ . Une translittération directe de ce résultat pour les courbes est faux.

En effet, soit  $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ . Alors  $\gamma'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$  et, en particulier,  $\|\gamma'(t)\| \equiv 1$ . En appliquant (fraudemment) le “théorème des valeurs intermédiaires” (théorème 3.1.6) à cette courbe paramétrée, disons sur l'intervalle  $[0, b] \subset [0, 2\pi]$ , cela voudrait dire qu'il existe  $s \in [0, b]$  tel que  $\begin{pmatrix} \cos b - 1 \\ \sin b \end{pmatrix} = b \begin{pmatrix} -\sin s \\ \cos s \end{pmatrix}$ . Cependant, en prenant la norme des deux vecteurs de chaque côté de l'égalité, il apparaîtrait que  $\sqrt{2 - 2\cos b} = b$  (et ce indépendamment de  $s$ ). Cette dernière égalité impliquerait que  $b = 0$  (il faut un peu travailler pour voir que c'est vrai). Autrement dit, le seul cas où le résultat pourrait être vrai est si l'intervalle n'est qu'un point, c'est-à-d. c'est un cas complètement trivial.



Ainsi, pas d'analogues du théorème 3.1.6 des valeurs intermédiaires pour les courbes paramétrées (et conséquemment pour les applications en général). Ceci n'empêche pas qu'il est possible d'appliquer ce théorème à chaque fonction coordonnée (mais le  $s$  est alors différent pour chacune). Ceci mène à des complications qui sont essentiellement aussi embêtantes que celles qui se trouvent dans la démonstration ci-dessus (c'est le célèbre principe de conservation des emmerdements). ♠

**Exemple 4.3.5.** Une courbe paramétrée très importante est la cycloïde alias la brachistochrone alias la tautochrone. Il s'agit de  $\gamma : [0, \pi] \rightarrow \mathbb{R}^2$  définie par

$$\gamma(t) = \begin{pmatrix} t - \sin t \\ \cos t \end{pmatrix}.$$

Sa longueur est

$$\int_0^\pi \sqrt{(1 - \cos t)^2 + \sin^2 t} dt = \int_0^\pi \sqrt{2} \sqrt{1 - \cos t} dt = 4 \int_0^1 \frac{u}{\sqrt{1 - u^2}} du = 4$$

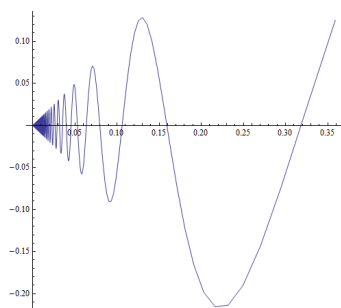
où le changement de variable  $\cos t = 1 - 2u^2$  a été utilisé. Il est bon de vérifier que cette longueur est plus grande que celle de la droite qui relie les deux points  $(\sqrt{\pi^2 + 2^2} \simeq 3,72)$ , car dans le plan, la droite est le chemin le plus court (comme il sera démontré à l'exercice 74). ♣

Si la courbe paramétrée est  $C^1$  par morceaux alors  $\|\gamma'(t)\|$  est bornée et continue en dehors d'un ensemble négligeable (des points). Par conséquent, la longueur est toujours définie. Par contre, en général, il existe des courbes  $C^0$  sur  $[0, 1]$ ,  $C^1$  sur  $]0, 1]$  et qui ne sont pas de longueur bornée.

**Exemple 4.3.6.** Soit  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  la courbe paramétrée définie par

$$\gamma(t) = \begin{pmatrix} t \\ t \sin \frac{\pi}{t} \end{pmatrix} \text{ si } t \neq 0 \text{ et } \gamma(0) = \mathbf{0}.$$

Alors  $\gamma'(t)$  est définie partout (sauf en 0) et  $\gamma(t)$  est continue, mais l'intégrale (impropre) de sa longueur ne converge pas (elle tend vers l'infini). Pour les détails, voir l'exercice 58. ♣



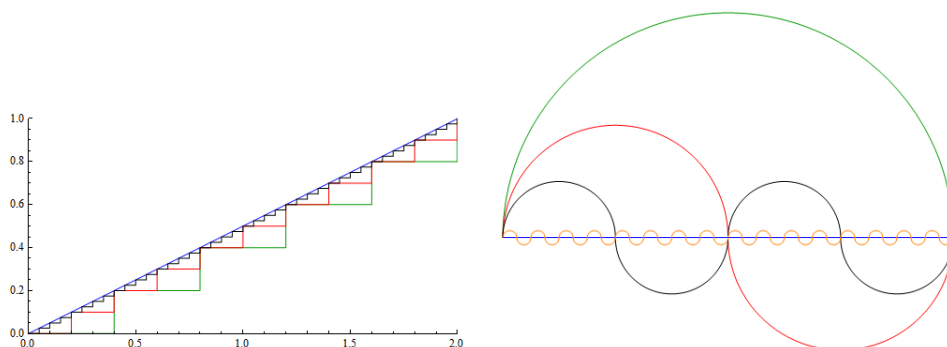
**EXERCICE 58:** Montrer que la courbe paramétrée  $\gamma(t) = \begin{pmatrix} t \\ t \sin \frac{\pi}{t} \end{pmatrix}$  définie pour  $t \in [0, 2]$  n'est pas rectifiable.

[Indice : trouver une suite de  $t_i$  telle que les sommes partielles des  $\|\gamma(t_{i+1}) - \gamma(t_i)\|$  se minorant par une somme partielle harmonique.]

**Remarque 4.3.7.** Néanmoins, lorsque  $I$  n'est pas un intervalle borné ou que  $\gamma'$  n'est pas bornée, il est possible de montrer que la proposition 4.3.2 reste vraie, i.e. que si l'intégrale impropre est divergente la courbe n'est pas rectifiable. ♠

**Remarque 4.3.8.** Une erreur qui fut fatale à plusieurs est de croire que la longueur d'une courbe est quelque chose qui se comporte bien, *c'ad.* que si deux courbes sont proches, alors leur longueurs sont proches. C'est totalement faux.

Un premier exemple est la suite de courbes qui consiste à approcher la diagonale d'un triangle rectangle par un escalier dont les marches sont de plus en plus fines. En effet, si  $a$  est la base,  $b$  la hauteur et  $c$  l'hypoténuse du triangle, alors  $c^2 = a^2 + b^2$  par Pythagore. Cependant, un escalier (aussi proche de l'hypoténuse soit-il) aura longueur  $a + b$  (voir dessin ci-dessous à gauche). Remarquez que  $a + b \geq \sqrt{a^2 + b^2}$ .



De la même manière, les points  $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$  et  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  de  $\mathbb{R}^2$  peuvent être reliés par un demi-cercle de rayon 1, ou encore deux demi-cercles de rayons  $\frac{1}{2}$  (centrés en  $\begin{pmatrix} -1/2 \\ 0 \end{pmatrix}$  et  $\begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$ ), ou quatre demi-cercles de rayons  $\frac{1}{4}$ , ou ... L'impression est que plus ces divisions sont faites, plus la courbe est proche du segment (voir dessin ci-dessus à droite). Pourtant sa longueur reste tout le temps égale à  $\pi$ . De nouveau  $\pi \geq 2$ .

Il est même possible de trouver des suites de courbes ( $C^1$  par morceaux) dont la longueur devient infinie qui sont très proches d'un segment (par exemple des fractales). En fait la longueur est, pour la topologie intuitive sur l'espace des courbes, semi-continues inférieurement, *c'ad.* que la longueur obtenue par un processus limite est toujours plus grande que la longueur de la courbe limite. ♠

### Proposition 4.3.9

Soit  $\gamma : I_1 \rightarrow \mathbb{R}^n$  une courbe paramétrée  $C^1$  par morceaux et  $\sigma : I_2 \rightarrow \mathbb{R}^n$  un reparamétrage  $C^1$  (i.e. paramétrage équivalent), alors  $\mathcal{L}ong(\sigma) = \mathcal{L}ong(\gamma)$ .

**DÉMONSTRATION:** Soit  $f : I_2 \rightarrow I_1$  un difféomorphisme de classe  $C^1$  tel que  $\sigma = \gamma \circ f$ . Alors,  $\sigma'(t) = \gamma'(f(t))f'(t)$  et en utilisant la formule de changement de variable :

$$\begin{aligned} \mathcal{L}ong(\gamma) &= \int_{I_1} \|\gamma'(t)\| dt = \int_{f(I_2)} \|\gamma'(t)\| dt \\ &= \int_{I_2} |f'(t)| \|\gamma'(f(t))\| dt = \int_{I_2} \|\gamma'(f(t))f'(t)\| dt \\ &= \int_{I_2} \|\sigma'(t)\| dt. \end{aligned}$$

Ce qui est la longueur de  $\sigma$ . ■

**EXERCICE 59:** La chaînette (ou caténaire, du latin catena signifiant chaîne) est la trace de la courbe paramétrée  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} t \\ \cosh t \end{pmatrix}$ . Paramétrer cette courbe par la longueur d'arc.

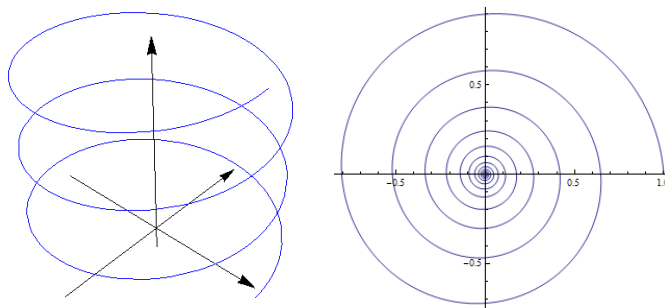
### 4.3 - Longueur ; cas $C^1$ par morceaux

[Quelques relations utiles sur les fonction hyperboliques :  $\operatorname{cht} = \frac{e^t + e^{-t}}{2}$ ,  $\operatorname{sh}t = \frac{e^t - e^{-t}}{2}$ ,  $\operatorname{ch}^2 t - \operatorname{sh}^2 t = 1$ ,  $\operatorname{ch}'t = \operatorname{sh}t$  et  $\operatorname{sh}'t = \operatorname{cht}$ . Les fonctions inverses s'appellent (en français)  $\operatorname{argsh}t = \ln(t + \sqrt{t^2 + 1})$  (définie  $\forall t \in \mathbb{R}$ ) et  $\operatorname{argcht} = \ln(t + \sqrt{t^2 - 1})$  (définie sur  $\mathbb{R}_{\geq 1}$ ).]

[Cette trace est la forme qu'aurait une corde (ou chaîne) inélastique et de masse uniforme suspendue entre deux points, lorsqu'elle se courbe sous son propre poids. En la regardant la tête à l'envers, cela donne ainsi la forme de l'arc qui sous l'effet de la gravité subit le moins de force possible (seule la force de sustentation). Il a longtemps été supposé que cette forme était celle d'une parabole. Suite à un défi de Jacques Bernoulli, les mathématiciens Leibniz, Jean Bernoulli et Huygens ont trouvé indépendamment la réponse en 1691. Beaucoup d'architectes s'en sont servi lorsque les matériaux plus modernes (e.g. le béton) l'ont rendu possible, e.g. Gaudi. Par ailleurs, un détour par Florence vous convaincra qu'une section du dôme de la cathédrale est plus proche d'une chaînette que d'une parabole.]

EXERCICE 60: Soit  $a, b \in \mathbb{R}$ . L'hélicoïde est la courbe paramétrée  $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^3$  définie par  $\gamma(t) = \begin{pmatrix} a \cos t \\ a \sin t \\ bt \end{pmatrix}$ . En trouver la longueur.

[Ci-dessous un hélicoïde (à gauche) et une spirale logarithmique (à droite).]



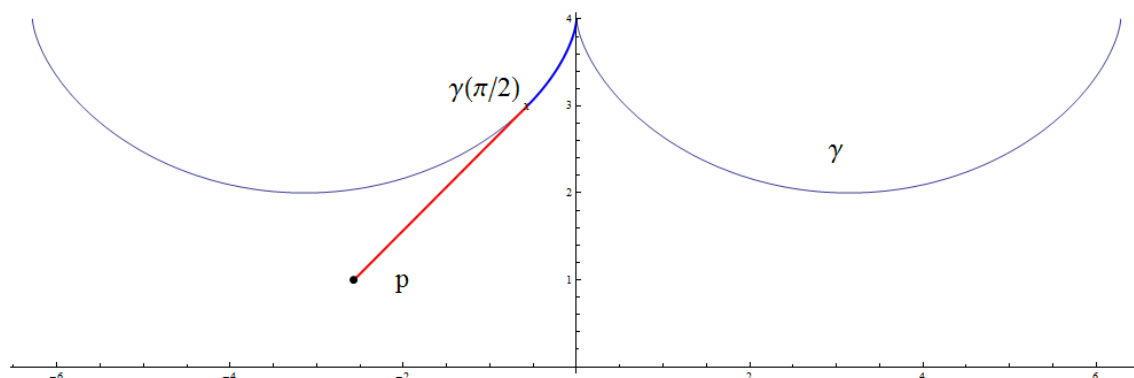
EXERCICE 61: La spirale logarithmique est la trace de la courbe paramétrée  $\gamma : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} e^{-t} \cos t \\ e^{-t} \sin t \end{pmatrix}$ .

- Trouver un autre paramétrage  $\sigma$ , tel que  $\sigma : [0, 1] \rightarrow \mathbb{R}^2$  a pour image  $\gamma(\mathbb{R}_{\geq 0}) \cup \mathbf{0}$ . [Indice : inverser l'application  $t \mapsto \frac{t}{1+t}$  qui envoie  $\mathbb{R}_{\geq 0}$  sur  $[0, 1[$ .]
- Montrer que ce paramétrage est  $C^1$ . [Seule une des deux extrémités de  $[0, 1]$  peut créer des soucis.]
- Pourquoi  $\sigma$  est-elle de longueur finie ?
- Trouver la longueur de cette trace. [Il est fortement recommandé d'utiliser plutôt le paramétrage  $\gamma$ , même si l'intégrale est impropre. Comme la courbe paramétrée n'est pas définie sur un intervalle, il faut regarder les longueurs de  $\gamma|_{[0, c]}$  pour  $c \in \mathbb{R}_{>0}$  puis faire tendre vers  $+\infty$ .]

EXERCICE 62: Le cycloïde/brachistochrone/tautochrone est la trace de la courbe paramétrée  $\gamma : [-2\pi, 2\pi] \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} t - \sin t \\ 3 + \cos t \end{pmatrix}$ . Imaginons qu'elle est réalisée par une structure rigide et qu'un fil (inextensible) de longueur 4 est suspendu en  $\gamma(0) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$  et qu'une masse est accrochée au bout du fil. Le but est de montrer que le pendule décrit un cycloïde sous oscillation. Soit  $t \in [-\pi, \pi]$ .



- Soit  $\gamma(t)$  le point où le fil quitte la structure métallique. Donner la longueur de la cycloïde entre  $\gamma(t)$  et  $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$ .
- Si  $t \neq 0$ , calculer la tangente en  $\gamma(t)$ .
- Le fil suivra la tangente à partir de  $\gamma(t)$ , donner  $\mathbf{p}$  la position de la masse. Montrer qu'il s'agit à translation près du cycloïde initial.



La courbe décrite par le pendule dans l'exercice précédent est une "développante" du cycloïde. L'exercice suivant montre l'importance de cette courbe en horlogerie.

**EXERCICE 63:** *Le but de cet exercice est de montrer que tout corps glissant sans frottement sur une cycloïde prend toujours le même temps pour atteindre le creux  $\gamma(\pi)$  (dans le paramétrage de l'exercice 62).*

- Calculer la longueur d'arc entre  $\gamma(\pi)$  et  $\gamma(t)$ . Cette "distance au creux" sera appelée dorénavant  $z$ .
- Au point  $\gamma(t)$ , l'accélération donnée par la gravité dans la direction de la courbe est  $\frac{\gamma'(t)}{\|\gamma'(t)\|} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ . En déduire que  $z'' = -\frac{1}{2}z$ .
- Résoudre cette équation (i.e. écrire  $z(\tau)$ ) en utilisant les conditions initiales  $z(0) = z_0$  et  $z'(0) = 0$ .
- Conclure.

## 4.4 Paramétrage par la longueur d'arc

Il est un paramétrage particulièrement utile pour les courbes paramétrées régulières. Il s'agit de l'abscisse curviligne, ou plus vulgairement, le paramétrage par la longueur d'arc.

**Définition 4.4.1.** Une courbe paramétrée est dite paramétrée par la longueur d'arc, si le vecteur vitesse est unitaire (lorsqu'il est défini), i.e.  $\|\gamma'(t)\| = 1$ . ★

Il est aussi possible de parler de "paramétrée par la longueur d'arc par morceaux", dans quel cas, en un nombre fini de point seul les limites à gauche et à droite de  $\gamma'(t)$  existent et sont de norme 1.

S'il existe, ce paramétrage n'est pas tout à fait unique, puisque le paramétrage dans le sens inverse a la même propriété. Par contre, si un paramétrage est équivalent, il faut que le difféomorphisme réalisant l'équivalence  $f$  satisfasse  $|f'(t)| = \pm 1$  ce qui implique (comme  $f$  est  $C^1$ ) que

$f'(t) \equiv 1$  ou  $-1$ . Ainsi, deux paramétrages par longueur d'arc équivalents sont reliés par un  $f$  de la forme  $t \mapsto a \pm t$  où  $a \in \mathbb{R}$ . Ceci est aussi relié au fait que les courbes dans  $\mathbb{R}^m$  n'ont que deux orientations (direction de parcours) possibles. Montrons maintenant que ce paramétrage existe dans une classe raisonnable de cas.

Une remarque simple, préalable à l'existence de tels paramétrages, est qu'une courbe régulière doit forcément avoir une longueur strictement positive (l'intégrale d'une quantité  $> 0$  sur un intervalle de longueur  $> 0$  est  $> 0$ ).

**Proposition 4.4.2**

Soit  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  une courbe paramétrée régulière, soit  $L = \text{Long}(\gamma) > 0$ , alors il existe  $g : [0, L] \rightarrow [a, b]$  un difféomorphisme  $C^1$  tel que  $\gamma \circ g$  est un paramétrage de  $\text{Im } \gamma$  par la longueur d'arc.

DÉMONSTRATION: La démonstration est essentiellement celle du théorème 4.2.4, mais où il y a plus de régularité. Soit  $\mathcal{A}(t) : [a, b] \rightarrow [0, L]$  définie par

$$\mathcal{A}(t) = \int_a^t \|\gamma'(t)\| dt.$$

Par le théorème fondamental du calcul intégral,  $\mathcal{A}'(t) = \|\gamma'(t)\|$  et cette dernière quantité est  $> 0$  par régularité de  $\gamma$ . Ainsi,  $\mathcal{A}$  est de classe  $C^1$  et, comme  $\mathcal{A}'(t) > 0$ , le théorème d'inversion locale implique que  $\mathcal{A}$  est un difféomorphisme  $C^1$ . Par conséquent, il existe un  $g : [0, L] \rightarrow [a, b]$  tel que  $\mathcal{A} \circ g$  est l'identité. De plus  $g(s)' = \frac{1}{\mathcal{A}'(g(s))}$ . Ainsi, par dérivation en chaîne  $\phi = \gamma \circ g$  est un paramétrage de vitesse constante égale à 1. ■

Au vu des définitions ci-haut, une courbe est régulière si et seulement si elle possède un paramétrage par la longueur d'arc. La proposition ci-dessus montre la partie  $\Rightarrow$ , pour voir  $\Leftarrow$  il suffit de voir que si elle possède un paramétrage par la longueur d'arc de classe  $C^1$ , comme ce paramétrage est de vitesse non-nulle en tout point, la courbe est régulière.

**Exemple 4.4.3.** La courbe paramétrée  $\gamma : [-1, 1] \rightarrow \mathbb{R}^2$  régulière par morceaux définie par  $\gamma(t) = \frac{1}{\sqrt{2}} \begin{pmatrix} t \\ |t| \end{pmatrix}$  est paramétrée par la longueur d'arc par morceaux. ♣

**Exemple 4.4.4.** Reprenons le cas de la cycloïde,  $\gamma(t) = \begin{pmatrix} t - \sin t \\ \cos t \end{pmatrix}$ . La fonction longueur d'arc est

$$\mathcal{A}(s) = \int_0^s \sqrt{(1 - \cos t)^2 + \sin^2 t} dt = \int_0^s \sqrt{2} \sqrt{1 - \cos t} dt = 4 \int_0^{\sin \frac{s}{2}} \frac{u}{\sqrt{1 - u^2}} du = 4(1 - \cos \frac{s}{2})$$

en utilisant : le changement de variable  $\cos t = 1 - 2u^2$ , que  $\frac{1 - \cos s}{2} = \sin^2 \frac{s}{2}$ , et que le cosinus et le sinus de  $\frac{s}{2}$  sont positifs sur  $[0, \pi]$ . Pour trouver la fonction  $g$ , il est ici possible d'inverser "à la main" : posons  $y = \mathcal{A}(s)$ , alors

$$y = 4(1 - \cos \frac{s}{2}) \Leftrightarrow s = 2 \arccos(1 - \frac{y}{4}).$$

Il apparaît alors que  $\mathcal{A}(2 \arccos(1 - \frac{y}{4})) = y$ . La fonction de reparamétrage recherchée est  $g(t) = 2 \arccos(1 - \frac{t}{4})$ . Le nouveau paramétrage est  $\sigma(t) = \gamma(g(t)) = \gamma(2 \arccos(1 - \frac{t}{4}))$ . Comme  $\sin \arccos B = \sqrt{1 - B^2}$ , il apparaît que

$$\gamma(g(t)) = \left( 2 \arccos(1 - \frac{t}{4}) - \frac{(4-t)\sqrt{t^2-8t}}{8}, \frac{(4-t)^2-8}{8} \right),$$

où  $t \in [0, 4]$ , est le paramétrage par longueur d'arc de la trace  $\text{Im} \gamma$ . ♣

**Exemple 4.4.5.** Soit  $\gamma: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^2$  la courbe paramétrée (une moitié de parabole semi-cubique) définie par  $\gamma(t) = (t^{3/3}, t^{2/2})$ . Pour trouver un [re]paramétrage par la longueur d'arc, il faut d'abord calculer la fonction de la longueur d'arc :  $\gamma(t) = (t^2, t^3)$ , d'où

$$\mathcal{A}(s) = \int_0^s \sqrt{t^2 + t^4} dt = \int_0^s t \sqrt{1 + t^2} dt = \frac{1}{3} (1 + s^2)^{3/2} - \frac{1}{3}.$$

De nouveau, ceci s'inverse facilement à la main :

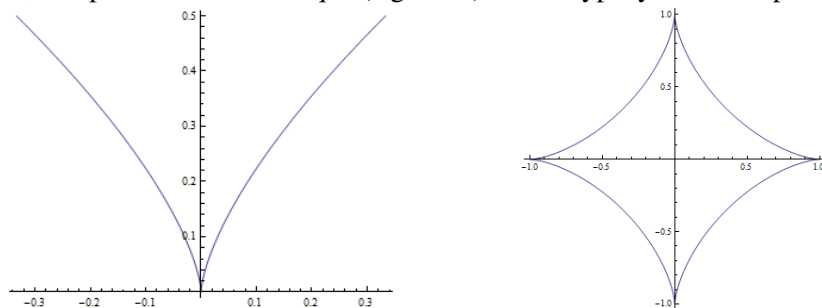
$$\mathcal{A} = \frac{1}{3} (1 + s^2)^{3/2} - \frac{1}{3} \Leftrightarrow \sqrt{(3\mathcal{A} + 1)^{2/3} - 1} = s.$$

Ainsi, la courbe paramétrée

$$\sigma(t) = \gamma\left(\sqrt{(3t+1)^{2/3} - 1}\right) = \left( \frac{1}{3} \left( (3t+1)^{2/3} - 1 \right)^{3/2}, \frac{(3t+1)^{2/3} - 1}{2} \right)$$

est un [re]paramétrage par la longueur d'arc de  $\gamma$ . ♣

[Ci-dessous, une parabole semi-cubique (à gauche) et une hypocycloïde à 4 pointes (à droite).]



EXERCICE 64: Trouver un paramétrage par la longueur d'arc (par morceaux) de la courbe paramétrée  $\gamma: [0, \frac{\pi}{2}] \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = (\cos^3 t, \sin^3 t)$  (il s'agit d'un morceau d'hypocycloïde [à 4 pointes]).

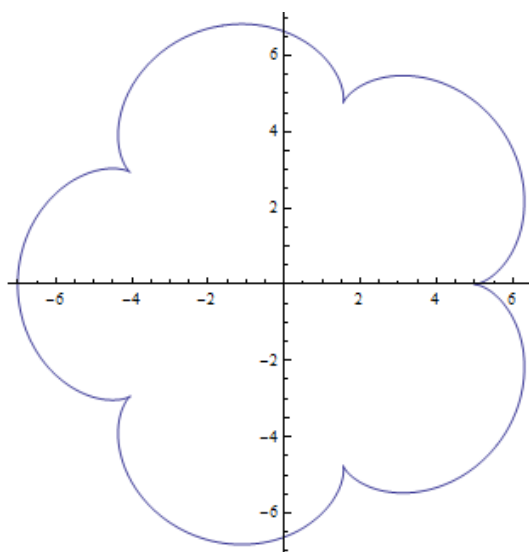
EXERCICE 65: Soit  $\gamma: [-\pi, \pi] \rightarrow \mathbb{R}^2$  la courbe paramétrée (nommée cardioïde) définie par  $\gamma(t) = (\cos t(1 + \cos t), \sin t(1 + \cos t))$ .

- Donner la longueur d'arc de  $\gamma$  au temps  $s \in [-\pi, \pi]$ , c'ad. calculer  $\mathcal{A}(s) = \text{Long}(\gamma|_{[0,s]})$ .
- Paramétrer la courbe  $\gamma$  par la longueur d'arc, i.e. trouver  $f: [0, \text{Long} \gamma] \rightarrow [a, b]$  tel que, pour  $s \in [0, \text{Long} \gamma]$ ,  $\text{Long}(\gamma \circ f|_{[0,s]}) = s$ .

EXERCICE 66: Soit  $\gamma$  la courbe paramétrée (épicycloïde) définie de  $[0, 2\pi]$  dans  $\mathbb{R}^2$  par

$$\gamma(t) = (R \cos t - \cos(Rt), R \sin t - \sin(Rt)),$$

où  $R \in \mathbb{R}_{>0}$ . Voici une représentation de son image, pour  $R = 6$  :



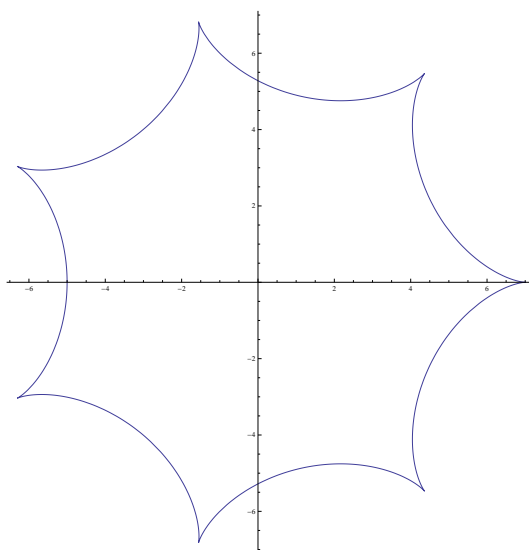
- En quels points ce paramétrage est-il de classe  $C^1$  ?
- Calculer le vecteur vitesse et sa norme. En quels points le paramétrage est-il régulier ?
- Trouver un paramétrage par la longueur d'arc (de classe  $C^1$  par morceaux mais pas  $C^1$ ). Déduire de ces calculs la longueur de la courbe.
- Lorsque  $R$  est entier,  $\gamma(0) = \gamma(2\pi)$ . Le champ de vecteur  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  défini par  $F(x, y) = \begin{pmatrix} -y \\ x \end{pmatrix}$ . Pourquoi le travail du champ de vecteur  $F$  le long de  $\gamma$  donne l'aire du domaine borné délimité par la courbe ?
- En déduire l'aire de ce domaine.

Voici l'exercice cousin des deux précédents :

EXERCICE 67: Soit  $\gamma$  la courbe paramétrée (hypocycloïde) définie de  $[0, 2\pi]$  dans  $\mathbb{R}^2$  par

$$\gamma(t) = (R \cos t + \cos(Rt), R \sin t - \sin(Rt)),$$

où  $R \in \mathbb{R}_{>0}$ . Voici une représentation de son image, pour  $R = 6$  :



- En quels points ce paramétrage est-il de classe  $C^1$  ?
- Calculer le vecteur vitesse et sa norme. En quels points le paramétrage est-il régulier ?
- Trouver un paramétrage par la longueur d'arc (de classe  $C^1$  par morceaux mais pas  $C^1$ ). Déduire de ces calculs la longueur de la courbe.
- Lorsque  $R$  est entier,  $\gamma(0) = \gamma(2\pi)$ . Le champ de vecteur  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  défini par  $\mathbf{F}(x, y) = \begin{pmatrix} -y \\ x \end{pmatrix}$ . Pourquoi le travail du champ de vecteur  $\mathbf{F}$  le long de  $\gamma$  donne l'aire du domaine borné délimité par la courbe ?
- En déduire l'aire de ce domaine.

EXERCICE 68: La développante du cercle est la courbe décrite par la main de quelqu'un qui déroule une bobine circulaire (fixe). Pour un cercle de rayon 1 centré en  $\mathbf{0} \in \mathbb{R}^2$ , ceci donne la courbe paramétrée  $\gamma : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^2$  définie par

$$\gamma(t) = \begin{pmatrix} \cos t + t \sin t \\ \sin t - t \cos t \end{pmatrix}.$$

Donner la fonction de longueur d'arc  $\mathcal{A} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  qui à  $t \in \mathbb{R}_{\geq 0}$  associe la longueur de la courbe entre  $\gamma(0)$  et  $\gamma(t)$ . Donner  $\sigma$  un reparamétrage  $C^1$  de  $\gamma$  qui est un paramétrage par la longueur d'arc.

EXERCICE 69: Xorqui a trouvé une courbe paramétrée fantastique qu'il décide de nommer le iou-  
teurnoïde rectibaladant. En se penchant sur sa feuille, vous y lisez le paramétrage  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  défini par

$$\gamma(t) = \begin{pmatrix} \ln \cosh t \\ 2 \operatorname{arctg} \left( \tanh \frac{t}{2} \right) \end{pmatrix}.$$

Donner la fonction de longueur d'arc  $\mathcal{A} : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  qui à  $t \in \mathbb{R}_{\geq 0}$  associe la longueur de la courbe entre  $\gamma(0)$  et  $\gamma(t)$ . Ce paramétrage a-t-il quelque chose de particulier ?

EXERCICE 70: Xorqui a trouvé une seconde courbe paramétrée fantastique qu'il décide de nommer la patatoïde parachevée. En se penchant sur sa feuille vous y lisez le paramétrage  $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$  défini par

$$\gamma(t) = \begin{pmatrix} 4 \cos t - \cos(4t) \\ 4 \sin t - \sin(4t) \end{pmatrix}.$$

- Calculer la norme du vecteur vitesse  $\gamma'$  de ce paramétrage. Trouver les points où le paramétrage n'est pas régulier (i.e. les points où le vecteur vitesse est nul).
- Donner la fonction de longueur d'arc  $\mathcal{A} : [0, 2\pi] \rightarrow \mathbb{R}_{\geq 0}$  qui à  $s \in [0, 2\pi]$  associe la longueur de la courbe entre  $\gamma(0)$  et  $\gamma(s)$ . [Attention :  $\sqrt{a^2} = |a|$ .]
- À partir de la fonction  $\mathcal{A}$  trouvée, indiquer pourquoi il n'y a pas de reparamétrage  $C^1$  de  $\gamma$  qui est un paramétrage par la longueur d'arc. [Attention : Le problème est lié à la régularité ; il est possible de trouver un reparamétrage  $C^0$  qui donne, en les points où la vitesse est définie, un paramétrage par la longueur d'arc.]

EXERCICE 71: Soit  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$  la courbe définie par  $\gamma(t) = \begin{pmatrix} t^4/4 \\ t^6/6 \end{pmatrix}$ .

- Est-ce un paramétrage régulier ?

b. Soit  $\sigma$  le paramétrage de  $\gamma$  restreint à  $[0, 1]$ , i.e.

$$\begin{aligned} \sigma: [0, 1] &\rightarrow \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} t^4/4 \\ t^6/6 \end{pmatrix} \end{aligned}$$

Quelle est la longueur de la trace de  $\sigma$  ?

c. Trouver un reparamétrage par la longueur d'arc de  $\sigma$ .

## 4.5 Intégrale curviligne et travail

Il existe deux notions d'intégration le long d'une courbe, la première qui est le plus souvent utilisée en mathématiques (cf. l'analyse complexe) est l'intégrale curviligne, et la seconde qui est plutôt utilisée en physique est la notion de travail.

Étant donné une fonction  $F: U \rightarrow \mathbb{R}$ , où  $U$  est un ouvert de  $\mathbb{R}^m$ , et une courbe paramétrée  $\gamma: [a, b] \rightarrow \mathbb{R}^m$  de classe  $C^1$  telle que  $\gamma([a, b]) \subset U$ , il est tentant de penser à l'intégrale de  $F$  le long de la trace de  $\gamma$  comme à l'intégrale de  $F \circ \gamma$ . Cette dernière dépend malheureusement du paramétrage choisi. Or, il serait souhaitable que la valeur de l'intégrale de la fonction constante 1 sur un domaine est le volume du domaine. En particulier, que la valeur ne dépende pas du paramétrage. Comme ici le domaine est une trace (l'image d'une courbe paramétrée), il serait naturel de retrouver la longueur de cette trace. Pour ces deux raisons, la définition suivante s'impose.

**Définition 4.5.1.** L'intégrale curviligne d'une fonction  $F: U \rightarrow \mathbb{R}$ , où  $U$  est un ouvert de  $\mathbb{R}^m$ , le long d'une courbe paramétrée  $\gamma: [a, b] \rightarrow U$  de classe  $C^1$  par morceaux est

$$\int_{\gamma} F = \int_a^b F(\gamma(t)) \|\gamma'(t)\| dt \quad \star$$

Comme voulu,  $\mathcal{L}ong(\gamma) = \int_{\gamma} 1$  et, avec les mêmes arguments que pour montrer l'invariance de la longueur sous paramétrages équivalents, il est possible de vérifier que cette quantité ne dépend pas de la classe d'équivalence du paramétrage choisi.

EXERCICE 72: Montrer que la définition d'intégral curviligne est invariante sous reparamétrage  $C^1$ .

Le centre de gravité d'une courbe  $\gamma: [a, b] \rightarrow \mathbb{R}^m$  est le centre de masse d'un fil (rigide, e.g. de cuivre) de densité constante. Formellement, si  $\gamma(t) = (\gamma_1(t), \gamma_2(t), \dots, \gamma_m(t))^T$ , c'est le vecteur  $\mathbf{x}$  tel que

$$\pi_i(\mathbf{x}) = x_i = \frac{1}{\mathcal{L}ong(\gamma)} \int_{\gamma} \pi_i = \frac{1}{\mathcal{L}ong(\gamma)} \int_a^b \pi_i(\gamma(t)) \|\gamma'(t)\| dt = \frac{1}{\mathcal{L}ong(\gamma)} \int_a^b \gamma_i(t) \|\gamma'(t)\| dt.$$

**Exemple 4.5.2.** Soit  $\gamma: [-1, 1] \rightarrow \mathbb{R}^2$  la courbe paramétrée définie par  $\gamma(t) = (t, \sqrt{1-t^2})$ . Sa trace est un demi-cercle et  $\|\gamma'(t)\| = \frac{1}{\sqrt{1-t^2}}$ . Sa longueur se calcule (ou se vérifie, ou se devine) comme étant  $\pi$ . Son centre de gravité est ainsi

$$\frac{1}{\pi} \begin{pmatrix} \int_{-1}^1 \frac{t}{\sqrt{1-t^2}} dt \\ \int_{-1}^1 1 dt \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{\pi} \end{pmatrix}. \quad \clubsuit$$

Un autre concept important où intervient l'intégration le long d'une courbe en physique est le travail. Rappelons que, dans un cas simple, si une force  $\mathbf{F}$  est exercée dans le même sens qu'un déplacement  $s$  (linéaire), le travail est alors  $Fs$ . Lorsque la force est dans le sens contraire du déplacement, la convention est alors que le travail est  $-Fs$ . Et quand il n'y a pas colinéarité, il suffit de regarder la composante de la force qui est dans le sens du mouvement.

Dans un cas plus compliqué, une particule se déplace dans  $\mathbb{R}^m$  et son mouvement est décrit par une courbe paramétrée  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  (c'ad. qu'au temps  $t \in [a, b]$ , la particule se trouve en  $\gamma(t) \in \mathbb{R}^m$ ). Supposons qu'une des forces qui s'opère sur elle ne dépend que de sa position. Autrement dit, la force est donnée par un champ de vecteurs  $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ . Il est alors possible de généraliser la notion ci-dessus du travail qu'effectue la force  $\mathbf{F}$  le long de  $\gamma$  :

**Définition 4.5.3.** Soit  $U \subset \mathbb{R}^m$  un ouvert, soit  $\gamma : [a, b] \rightarrow U$  une courbe  $C^1$  par morceaux et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  un champ de vecteurs. Alors le travail de  $\mathbf{F}$  le long de  $\gamma$  est, lorsque définie, l'intégrale

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{s} = \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt. \quad \star$$

Le travail n'est pas tout à fait indépendant du paramétrage. En effet, l'expression  $\mathbf{F}(\gamma(t)) \cdot \gamma'(t)$  s'écrit aussi  $\|\mathbf{F}(\gamma(t))\| \|\gamma'(t)\| \cos \angle(\mathbf{F}(\gamma(t)), \gamma'(t))$ . Si le cos ne dépendait pas de  $\gamma'(t)$ , le contexte serait de nouveau celui de l'intégrale curviligne. Cependant, deux paramétrages équivalents ont en un point fixé de la trace des vecteurs tangents colinéaires, ainsi le cos ne peut au plus changer que d'un signe. Comme un difféomorphisme  $f$  entre deux intervalles est soit croissant soit décroissant, le signe de  $f'$  donnera cette différence. D'où :

**Proposition 4.5.4**

Soit  $U$  un ouvert, soit  $\gamma_1 : [a, b] \rightarrow U$  une courbe paramétrée de classe  $C^1$  par morceaux, soit  $f : I \rightarrow [a, b]$  un difféomorphisme de classe  $C^1$ , et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  un champ de vecteurs. Si  $\int_{\gamma_1} \mathbf{F} \cdot d\mathbf{s}$  existe et  $\gamma_2 = \gamma_1 \circ f$ , alors  $\int_{\gamma_2} \mathbf{F} \cdot d\mathbf{s} = \text{sgn}(f') \int_{\gamma_1} \mathbf{F} \cdot d\mathbf{s}$ .

EXERCICE 73: Montrer la proposition ci-dessus.

**Exemple 4.5.5.** Soit  $\mathbf{F}(\mathbf{x}) = \mathbb{R}^m \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^m$  le champ de force défini par  $\mathbf{F}(\mathbf{x}) = \frac{c\mathbf{x}}{\|\mathbf{x}\|^3}$  (e.g. le champ de forces gravitationnelles ou électriques engendré par une masse ou une charge). Soit  $\mathbf{z} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ , soit  $a \in \mathbb{R}$  et soit  $\gamma_1 : [0, a] \rightarrow \mathbb{R}^m \setminus \{\mathbf{0}\}$  la courbe paramétrée définie par  $\gamma_1(t) = (1+t)\mathbf{z}$ . Alors  $\gamma_1'(t) = \mathbf{z}$ , et

$$\int_{\gamma_1} \mathbf{F} \cdot d\mathbf{s} = \int_0^a \frac{c\mathbf{z}}{(1+t)^2 \|\mathbf{z}\|^3} \cdot \mathbf{z} dt = \frac{c}{\|\mathbf{z}\|} \int_0^a \frac{1}{(1+t)^2} dt = \frac{c}{\|\mathbf{z}\|} \left(1 - \frac{1}{1+a}\right).$$

Soit maintenant  $\gamma_2 : [a, b] \rightarrow \mathbb{R}^m \setminus \{\mathbf{0}\}$  définie par  $\gamma_2(t) = \|\mathbf{z}\|(\cos t, \sin t)$ . Alors  $\gamma_2'(t) = \|\mathbf{z}\|(-\sin t, \cos t)$  et,  $\forall t \in [a, b]$ ,  $\mathbf{F}(\gamma_2(t)) \cdot \gamma_2'(t) = 0$ . Ainsi, le travail de  $\mathbf{F}$  le long de  $\gamma_2$  est nul.

Il est possible de construire, en combinant (plusieurs fois) ces deux chemins, toutes sortes de chemins  $C^1$  par morceaux allant de  $\mathbf{x}$  à  $\mathbf{y}$  tels que le travail de  $\mathbf{F}$  le long de ces chemins est (indépendamment du chemin)  $\frac{c}{\|\mathbf{x}\|} - \frac{c}{\|\mathbf{y}\|}$ . Ceci illustre une propriété des champs vectoriels qui découlent d'un potentiel ; une idée qui sera explorée plus tard. ♣

EXERCICE 74: “Les droites minimisent la longueur.” Soit  $\gamma: [a, b] \rightarrow \mathbb{R}^m$  une courbe paramétrée telle que  $\gamma(a) = \mathbf{p}$  et  $\gamma(b) = \mathbf{q}$ . Montrer que pour tout vecteur  $\mathbf{v}$  de norme  $\|\mathbf{v}\| = 1$

$$(\mathbf{q} - \mathbf{p}) \cdot \mathbf{v} = \int_a^b \gamma'(t) \cdot \mathbf{v} dt \leq \int_a^b \|\gamma'(t)\| dt.$$

Dans quel cas y a-t-il égalité ? Ensuite montrer que

$$\|\gamma(a) - \gamma(b)\| \leq \int_a^b \|\gamma'(t)\| dt,$$

c’ad. que la courbe la plus courte de  $\mathbf{p}$  à  $\mathbf{q}$  est le segment de droite joignant les deux points. Déterminer toutes les courbes (de classe  $C^1$  par morceaux) qui réalisent l’égalité.

EXERCICE 75: Soit  $\gamma: I \rightarrow \mathbb{R}^m$  une courbe régulière. Soit  $\mathbf{x}$  un point qui n’est pas sur la trace, et soit  $\mathbf{y} = \gamma(t_0)$  le point sur la courbe qui est plus le proche de  $\mathbf{x}$  (il n’est pas forcément unique), c’ad. :

$$\forall t \in I, \|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \gamma(t)\|.$$

Dire pourquoi il existe un tel point  $\mathbf{y}$ . Montrer que le vecteur  $\mathbf{x} - \mathbf{y}$  est orthogonal à la courbe au point  $\mathbf{y}$ .

## 4.6 Courbure

Même si pour l’instant seule la vitesse a été définie, il est (d’un point de vue de la physique newtonienne) normal de parler d’accélération. Le vecteur accélération est simplement défini comme étant  $\gamma''(s) = \frac{d}{ds}\gamma'(s) = \frac{d^2}{ds^2}\gamma(s) = D_1^2\gamma(s)$ .

Pour analyser plus finement le comportement des courbes il faut introduire un triplet de vecteurs unitaires. Le premier d’entre eux est le vecteur unitaire tangent,

$$\hat{\mathbf{T}}_\gamma(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|}.$$

Lorsqu’une courbe est paramétrée par la longueur d’arc,  $\gamma'(s) = \hat{\mathbf{T}}_\gamma(s)$ . Il est possible de voir  $\hat{\mathbf{T}}_\gamma$  comme une nouvelle courbe paramétrée, elle aura comme particularité que sa trace ( $\text{Im } \hat{\mathbf{T}}_\gamma$ ) est dans la sphère de  $\mathbb{R}^m$  (l’ensemble des points de norme 1). Le lemme suivant donne une propriété utile de ces courbes.

### Lemme 4.6.1

Soit  $\sigma: [a, b] \rightarrow \mathbb{R}^m$  une courbe paramétrée de classe  $C^1$  telle que  $\forall t \in [a, b], \|\sigma(t)\| = 1$ . Alors  $\sigma'(t)$  est orthogonal à  $\sigma(t)$ , i.e.  $\forall t \in [a, b], \sigma'(t) \cdot \sigma(t) = 0$ .

DÉMONSTRATION: Tout d’abord,  $\|\sigma(s)\|^2 = \sigma(s) \cdot \sigma(s) = 1$ . En dérivant la dernière égalité, il s’obtient que

$$\sigma'(s) \cdot \sigma(s) + \sigma(s) \cdot \sigma'(s) = 0 \Leftrightarrow \sigma'(s) \cdot \sigma(s) = 0. \quad \blacksquare$$

En particulier, lorsque la courbe est paramétrée par la longueur d’arc, le vecteur accélération est orthogonal au vecteur vitesse. En fait, pour un tel paramétrage, la variation du vecteur unitaire tangent exprime à quelle vitesse la courbe tourne. En effet, voici le cas modèle :



**Exemple 4.6.2.** Soit  $r \in \mathbb{R}_{>0}$  et soit  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$  la courbe paramétrée définie par  $\gamma(t) = \begin{pmatrix} r \cos(t/r) \\ r \sin(t/r) \end{pmatrix}$  (la trace de cette courbe est le cercle). Ce paramétrage est déjà un paramétrage par la longueur d'arc et

$$\gamma'(t) = \begin{pmatrix} -\sin \frac{t}{r} \\ \cos \frac{t}{r} \end{pmatrix} \quad \text{et} \quad \gamma''(t) = -\frac{1}{r} \begin{pmatrix} \cos \frac{t}{r} \\ \sin \frac{t}{r} \end{pmatrix}.$$

En particulier,  $\|\gamma''(t)\| = r^{-1}$ , *i.e.* l'accélération est plus grande quand le rayon est petit. Ceci est justifié comme suit, comme toute courbe “parcourt” une distance  $t$  en un temps  $t$ , plus le rayon est petit, plus la courbe fait un tour rapidement et donc plus le vecteur tangent unitaire fait (lui aussi) un tour rapidement. ♣

Ceci est à la base de la définition suivante.

**Définition 4.6.3.** La courbure d'une courbe paramétrée  $\gamma$  paramétrée par la longueur d'arc est définie par  $\kappa_\gamma(t) = \|\frac{d}{dt} \hat{T}_\gamma(t)\| = \|\gamma''(t)\|$ . Le vecteur unitaire normal est le vecteur  $\hat{N}_\gamma(t)$  tel que  $\kappa_\gamma(t) \hat{N}_\gamma(t) = \gamma''(t)$ . ★

Par le lemme 4.6.1, le vecteur normal unitaire  $\hat{N}$  est orthogonal au vecteur tangent unitaire  $\hat{T}$ . Le plan engendré par le vecteur unitaire tangent et le vecteur unitaire normal est appelé plan osculateur.

**Remarque 4.6.4.** D'autre part, la physique nous apprend deux choses : a- la force est proportionnelle à l'accélération et b- la force nécessaire pour qu'une particule à vitesse constante décrivent un cercle est inversement proportionnelle au rayon de ce cercle. Ceci n'est qu'une reformulation de ce qui est écrit plus haut. Par contre, cette remarque est intéressante pour ce qui est du tracé de chemins de fers (et aussi de routes). En effet, un train suivant son trajet exercera une force latérale sur les rails dès qu'il y a courbure. Un aspect important lors du tracé est de faire en sorte que la courbure n'augmente jamais de manière brutale. Les mauvais tracés (faits par exemple en mettant bout à bout des arcs de cercles et des droites) ont pour effet d'user prématurément les rails (voire de provoquer des accidents).

De la même manière, pour les tracés de route, de telles combinaisons (droites et cercles) sont à éviter pour que le conducteur n'ait pas à tourner brusquement le volant [et/ou que les passagers ne soient pas trop bringueballés]. ♠

Plusieurs courbes (en particulier, la lemniscate et la spirale de Cornu [alias clothoïde]) voient ainsi leur apparition sur les cartes et les vues du ciel. Par exemple, dans une lemniscate, la courbure est proportionnelle à la distance à l'origine. En parcourant une spirale de Cornu à vitesse constante, la courbure croît de façon linéaire (dans une voiture à vitesse constante il faudrait tourner le volant de manière progressive).

La courbure a une interprétation géométrique simple comme le rayon d'un cercle qui est tangent à la courbe à ce point (son centre est dans la direction de  $\hat{N}$ ). Si la courbe est plane en cet endroit, tout cercle de rayon plus grand (et inscrit dans le plan osculateur) passant par  $\gamma(t)$  devrait couper la courbe. De ce fait, le rayon de ce cercle en  $\gamma(t)$  est le rayon de courbure de la courbe  $\gamma$  en  $\gamma(t)$  est défini comme étant  $R_\gamma(t) = \frac{1}{\kappa_\gamma(t)}$  (dans le cas de 4.6.2, ceci est bien le rayon du cercle décrit).

**Définition 4.6.5.** Le centre de courbure au point  $\gamma(t)$  est le centre de ce cercle, autrement dit,  $\sigma(t) = \gamma(t) + \frac{\hat{N}_\gamma(t)}{\kappa_\gamma(t)}$ .

La nouvelle courbe paramétrée  $\sigma$  ainsi obtenue est appelée la courbe développée de  $\gamma$ . ★

Lorsque la courbe n'est pas paramétrée par la longueur d'arc, il est possible de calculer la courbure comme suit :

$$\kappa_\gamma(t) = \frac{\sqrt{\|\gamma'(t)\|^2 \|\gamma''(t)\|^2 - (\gamma'(t) \cdot \gamma''(t))^2}}{\|\gamma'(t)\|^3}$$

$$\text{et } \hat{N}_\gamma(t) = \frac{\|\gamma'(t)\|^2 \gamma''(t) - (\gamma'(t) \cdot \gamma''(t)) \gamma'(t)}{\|\gamma'(t)\| \sqrt{\|\gamma'(t)\|^2 \|\gamma''(t)\|^2 - (\gamma'(t) \cdot \gamma''(t))^2}}$$

Dans  $\mathbb{R}^3$ , ces formules se simplifient :

$$\kappa_\gamma(t) = \frac{\|\gamma'(t) \wedge \gamma''(t)\|}{\|\gamma'(t)\|^3} \quad \text{et} \quad \hat{N}_\gamma(t) = \frac{(\gamma'(t) \wedge \gamma''(t)) \wedge \gamma'(t)}{\|\gamma'(t) \wedge \gamma''(t)\| \|\gamma'(t)\|}$$

EXERCICE 76: Une fonction  $f : [a, b] \rightarrow \mathbb{R}$  peut-être vue comme une courbe paramétrée  $\gamma : [a, b] \rightarrow \mathbb{R}^2$  en posant  $\gamma(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix}$ .

- Donner (à l'aide d'une intégrale définie, mais qui ne peut être évaluée) un reparamétrage de  $\gamma$  par la longueur d'arc.
- Utiliser ce reparamétrage pour déterminer la courbure de  $\gamma$  en fonction de  $f$  et ses dérivées (vérifier votre réponse grâce à la formule ci-dessus).

EXERCICE 77: Xorqui dans son travail doit relier deux bouts de chemin de fer ensemble de manière à éviter que la transition soit brutale (cf. remarque 4.6.4). Pour simplifier, supposons que le premier bout est la demi-droite  $\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 0 \text{ et } x \leq 0 \}$  et le second la demi-droite  $\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x \geq 2 \text{ et } 2y = x \}$ . Trouver une fonction polynomiale de degré le plus petit possible  $f : [0, 2] \rightarrow \mathbb{R}$  telle que l'image de  $\gamma : [0, 2] \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix}$  relie les deux bouts des demi-droites de manière à ce que le résultat soit une courbe lisse dont la courbure est nulle aux points de jonctions.

[Indice : il s'agit de conditions sur  $f$  et ses dérivées (jusqu'au second ordre) en  $x = 0$  et  $x = 2$ .]

L'exercice suivant a pour but de montrer que la courbure ne nécessite pas réellement que la courbe soit de classe  $C^2$ , c'ad. il y a une façon de définir la courbure pour les courbes de classe  $C^1$  tant que le vecteur tangent unitaire forme une courbe rectifiable.

EXERCICE 78: Soit  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  une courbe régulière et soit  $\phi : [a, b] \rightarrow \mathbb{R}^m$  la courbe paramétrée qui est donnée par le vecteur tangent unitaire en  $t$ , i.e.

$$\phi(t) := \frac{\gamma'(t)}{\|\gamma'(t)\|}$$

Notons  $I_\varepsilon$  l'intervalle  $[t, t + \varepsilon]$ , où  $\varepsilon \in \mathbb{R}_{>0}$ .

- Si  $\gamma$  est paramétrée par la longueur d'arc, montrer que

$$\kappa_\gamma(t) = \lim_{\varepsilon \rightarrow 0} \frac{\text{Long}(\phi|_{I_\varepsilon})}{\text{Long}(\gamma|_{I_\varepsilon})}$$

où les  $|_{I_\varepsilon}$  indiquent la restriction de l'application à l'intervalle  $I_\varepsilon$  et  $\kappa_\gamma(t)$  est la courbure de  $\gamma$  en  $t$ .

(b) Montrer que ceci reste vrai si  $\gamma$  n'est plus paramétrée par la longueur d'arc.

EXERCICE 79: Calculer la développée du cycloïde  $\gamma(t) = \begin{pmatrix} t - \sin t \\ \cos t \end{pmatrix}$ . Montrer que cette courbe est à nouveau une cycloïde (à une translation près de l'image et un reparamétrage simple).

EXERCICE 80: Calculer la développée de la tractrice  $\gamma(t) = \begin{pmatrix} \cos t \\ -\sin t + \log \tan \frac{t}{2} \end{pmatrix}$ . [Il s'agit d'une chaînette.]

Le but de l'exercice suivant est de trouver la formule de la courbure pour une courbe qui n'est pas paramétrée par la longueur d'arc.

EXERCICE 81: Soit  $\gamma$  une courbe régulière,  $\sigma = \gamma \circ f$  son reparamétrage par la longueur d'arc et  $f = \mathcal{A}^{-1}$ .

a. Montrer que  $f'(t) = 1/\|\gamma'(f(t))\|$ .

b. Montrer que  $f''(t) = -\frac{\gamma'(f(t)) \cdot \gamma''(f(t))}{\|\gamma'(f(t))\|^4}$ .

c. Montrer que  $\|\sigma''(t)\|^2 = \frac{1}{\|\gamma'(r)\|^6} \left( \|\gamma'(r)\|^2 \|\gamma''(r)\|^2 - (\gamma'(r) \cdot \gamma''(r))^2 \right)$  où  $r = f(t)$ .

d. Utiliser  $\|\mathbf{x} \wedge \mathbf{y}\|^2 + (\mathbf{x} \cdot \mathbf{y})^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$ , pour montrer que

$$\kappa_\gamma(r) = \frac{\|\gamma'(r) \wedge \gamma''(r)\|}{\|\gamma'(r)\|^3}$$

## 4.A Torsion

Lorsque  $m = 3$  il devient naturel de compléter la paire  $\hat{\mathbf{T}}, \hat{\mathbf{N}}$  pour en faire une base orthonormée, le troisième vecteur qui s'impose est alors le vecteur unitaire binormal,  $\hat{\mathbf{B}}_\gamma(t) = \hat{\mathbf{T}}_\gamma(t) \wedge \hat{\mathbf{N}}_\gamma(t)$ . Lorsque  $\hat{\mathbf{B}}$  reste constant, la courbe reste toujours dans le même plan (osculateur) où elle ne fait pas d'autre chose que de "courber" (d'avoir de la courbure). Si ce n'est pas le cas, la courbe est alors tordue (c'àd. elle a de la torsion), et ceci se manifeste par le fait que  $\hat{\mathbf{B}}'$  n'est pas trivial. Or,

$$\hat{\mathbf{B}}'(t) = \hat{\mathbf{T}}'(t) \wedge \hat{\mathbf{N}}(t) + \hat{\mathbf{T}}(t) \wedge \hat{\mathbf{N}}'(t) = \hat{\mathbf{T}}(t) \wedge \hat{\mathbf{N}}'(t),$$

car  $\hat{\mathbf{T}}'$  est proportionnel à  $\hat{\mathbf{N}}$ . Mais, par le lemme 4.6.1,  $\hat{\mathbf{N}}'$  est orthogonal à  $\hat{\mathbf{N}}$ . Ainsi  $\hat{\mathbf{B}}'$  est proportionnel à  $\hat{\mathbf{N}}$ . La torsion d'une courbe paramétrée  $\gamma$  est alors définie par le réel  $\tau_\gamma(t)$  tel que  $\hat{\mathbf{B}}'_\gamma(t) = \tau_\gamma(t) \hat{\mathbf{N}}_\gamma(t)$ .

Étant donné que  $\hat{\mathbf{T}}'$  et  $\hat{\mathbf{B}}'$  ont donné lieu à des quantités géométriques (la courbure et la torsion) il serait tentant de penser qu'il en serait de même avec la dérivée du vecteur normal  $\hat{\mathbf{N}}'$ . Cependant,  $\hat{\mathbf{N}}'_\gamma(t) = \hat{\mathbf{B}}_\gamma(t) \wedge \hat{\mathbf{T}}'_\gamma(t)$ , d'où

$$\begin{aligned} \hat{\mathbf{N}}'(t) &= \hat{\mathbf{B}}'(t) \wedge \hat{\mathbf{T}}(t) + \hat{\mathbf{B}}(t) \wedge \hat{\mathbf{T}}'(t) \\ &= \tau(t) \hat{\mathbf{N}}(t) \wedge \hat{\mathbf{T}}(t) + \kappa(t) \hat{\mathbf{B}}(t) \wedge \hat{\mathbf{N}}(t) \\ &= -\tau(t) \hat{\mathbf{B}}(t) - \kappa(t) \hat{\mathbf{T}}(t). \end{aligned}$$

Ainsi en chaque point de la courbe, les constructions ci-dessus décrivent un repère (appelé repère de Frenet ou Frenet-Serret). En oubliant les évaluations en un point de l'intervalle et la référence à  $\gamma$ , les relations s'écrivent

$$\frac{d}{ds} \begin{pmatrix} \hat{\mathbf{T}} \\ \hat{\mathbf{N}} \\ \hat{\mathbf{B}} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & -\tau \\ 0 & \tau & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{T}} \\ \hat{\mathbf{N}} \\ \hat{\mathbf{B}} \end{pmatrix}$$

Essentiellement, une courbe (dans l'espace de dimension 3) est complètement caractérisée par sa normale et sa torsion, le théorème (qui ne sera pas démontré ici) important sur le sujet va comme suit :

**Théorème 4.A.1 (Théorème fondamental de la théorie locale des courbes)**

Étant donné deux fonctions  $\kappa(s) > 0$  et  $\tau(s)$  de classe  $C^1$  définies sur un intervalle  $[0, L]$ , il existe une courbe paramétrée par la longueur d'arc  $\gamma: [0, L] \rightarrow \mathbb{R}^3$  telle que  $\kappa_\gamma = \kappa$  et  $\tau_\gamma = \tau$ . Si  $\sigma$  est une autre courbe qui satisfait les mêmes conditions, alors ces deux courbes ne diffèrent que par une translation et une isométrie linéaire de  $\mathbb{R}^3$  (i.e. une matrice dans  $O(3)$ , i.e. une matrice  $M$  telle que  $M^T = M^{-1}$ ).

Puisque les courbes qui sont dans le plan n'ont pas de torsion, il s'ensuit que la courbure caractérise les courbes dans le plan.

Lorsque les courbes ne sont pas paramétrées par la longueur d'arc, il est parfois pratique d'utiliser d'autres expressions pour la courbure et la torsion. Nous les citons ici sans démonstration :

$$\hat{\mathbf{T}}(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|}, \quad \hat{\mathbf{B}}(t) = \frac{\gamma'(t) \wedge \gamma''(t)}{\|\gamma'(t) \wedge \gamma''(t)\|}, \quad \hat{\mathbf{N}}(t) = \hat{\mathbf{B}}(t) \wedge \hat{\mathbf{T}}(t), \quad \kappa(t) = \frac{\|\gamma'(t) \wedge \gamma''(t)\|}{\|\gamma'(t)\|^3},$$

et

$$\tau(t) = -\frac{(\gamma'(t) \wedge \gamma''(t)) \cdot \gamma'''(t)}{\|\gamma'(t) \wedge \gamma''(t)\|^2}.$$

## 4.B Théorème fondamental de l'algèbre

Cette section présente rapidement une preuve du théorème fondamental de l'algèbre : tout polynôme (à coefficients complexes) possède une racine (complexe). En passant, quelques propriétés globales des courbes sont abordées.

**Définition 4.B.1.** Une courbe paramétrée  $\gamma: [a, b] \rightarrow \mathbb{R}^m$  est dite fermée si  $\gamma(a) = \gamma(b)$ . Si de plus  $\gamma$  est de classe  $C^1$ , elle est dite fermée seulement s'il est aussi vrai que  $\gamma'(a) = \gamma'(b)$ .

Une courbe est dite simple si elle est injective et fermée. ★

Ainsi, lorsque  $\gamma: [a, b] \rightarrow \mathbb{R}^m$  est une courbe paramétrée de classe  $C^1$  fermée alors, l'application unitaire tangente  $\hat{\mathbf{T}}_\gamma$  définie sur  $[a, b]$  prend valeur dans  $S^{m-1}$ , où  $S^{m-1} = \{\mathbf{x} \in \mathbb{R}^m \mid \|\mathbf{x}\| = 1\}$  est la sphère de dimension  $m - 1$ .

En particulier, lorsque  $m = 2$ , l'application tangente prend valeur dans le cercle. Il est possible de définir l'indice d'une courbe simple de classe  $C^1$  comme étant le nombre de tours que l'application

tangente fait en faisant attention à l'ordre, *c'àd.* deux tours dans le sens anti-horaire et un tour dans le sens horaire compte comme un tour dans le sens anti-horaire. Soit donc  $\text{Ind}(\hat{\mathbf{T}}_\gamma)$  comme le nombre de tours anti-horaires moins le nombre de tours horaires de la courbe décrite par le vecteur unitaire tangent.

### **Théorème 4.B.2**

Soit  $\gamma : [a, b] \rightarrow \mathbb{R}^2$  une courbe paramétrée fermée de classe  $C^1$ . Si  $\gamma$  est simple alors  $\text{Ind}(\hat{\mathbf{T}}_\gamma) = \pm 1$ . Si  $\text{Ind}(\hat{\mathbf{T}}_\gamma) = \pm 1$  alors soit  $\gamma$  est simple, soit si  $t, s \in [a, b]$  sont tels que  $\gamma(t) = \gamma(s)$  alors  $\gamma'(t)$  est colinéaire à  $\gamma'(s)$ .

En effet, il peut arriver qu'une courbe ne soit pas simple mais d'indice  $\pm 1$ , *e.g.* en traçant un 8 de sorte que le point d'intersection est très verticale et que la partie droite est dessinée, puis la partie gauche.

Démontrons (avec quelques boîtes noires) le théorème fondamental de l'algèbre. Soit  $k \in \mathbb{Z}_{\geq 1}$  et  $P(z) = \sum_{i \in \bar{k}} a_i z^i$  un polynôme de degré  $k$  (*i.e.*  $a_k \neq 0$ ). Supposons que  $a_0 \neq 0$  sinon  $z = 0$  est une racine évidente. Alors pour  $r \in \mathbb{R}_{\geq 0}$ , soit  $\gamma_r : [0, 2\pi] \rightarrow \mathbb{C} \simeq \mathbb{R}^2$ , définie par  $\gamma_r(t) = P(re^{it}) / \|P(re^{it})\|$  est en fait une courbe paramétrée fermée à valeur dans le cercle. Pour  $r$  assez grand, le terme de plus grand degré domine, ainsi il existe  $r_0$  tel que  $\text{Ind}(\gamma_{r_0}) = k$  (car l'indice de  $t \mapsto e^{ikt}$  pour  $t \in [0, 2\pi]$  est  $k$ ).

Autrement dit, la courbe paramétrée fermée  $\sigma : [0, 2\pi] \rightarrow \mathbb{R}^2$  définie par  $\sigma(t) = P(re^{it})$  fait plus d'un tour autour de l'origine. Si  $r_0$  est suffisamment grand il est aussi possible de supposer que  $\sigma$  entoure  $a_0$  aussi.

Cependant  $P(0) = a_0 \neq 0$ . Ainsi, la surface paramétrée  $(r, t) \mapsto P(re^{it})$  (pour  $r \in [0, r_0]$  et  $t \in [0, 2\pi]$ ) doit couvrir tous les points entre  $a_0$  et la trace de  $\sigma$ , en particulier l'origine. Il y a donc bel et bien une solution au polynôme.



## Chapitre 5

# Approximation et étude I

*“Quand les mouettes suivent un chalutier,  
c’est qu’elles pensent qu’on va leur jeter des sardines.”*

-E. CANTONA

Ce chapitre a pour but premier d’approximer les fonctions un peu plus précisément que par l’application linéaire qu’est la dérivée, et d’“étudier” les fonctions, *c’ad.* de mettre en place (sans garantie de succès) des techniques pour identifier leur comportement (croissance, décroissance, inflexions, ...). Ces deux thèmes ne seront abordés que pour des fonctions.

Dans le cas du second, la raison derrière cette restriction provient du fait que la notion d’ordre dans  $\mathbb{R}^n$  n’est pas naturelle, et par conséquent prendre une valeur plus élevée ne l’est pas non plus. Pour le premier, la raison est moins stricte : l’écriture du cas général demande des notations revêches qui ne sont pas d’une lisibilité exemplaire. Par surcroît, le résultat s’obtient de toute façon en regardant les fonctions coordonnées de l’application.

Des techniques plus sophistiquées (mais qui sont loin d’être infaillibles) pour décider si un point critique (un point où le gradient s’annule) est un extremum seront discutées ; elles font intervenir des approximations plus fines que l’approximation linéaire.

### 5.1 Formule de Taylor

La dérivée est l’approximation d’une application  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  par une application linéaire, ou dans un langage légèrement capillotracté un[e famille de  $n$ ] polynôme[s] (en  $m$  variables) de degré 1. Ceci permet de faire des approximations numériques directes et de procéder à l’étude de fonction. La formule de Taylor est essentiellement une tentative de faire une meilleure approximation. Celle-ci permettra d’améliorer à la fois les approximations numériques (et même d’en produire pour des fonctions qui sont définies abstraitement, voir section 6.2) et la détection d’extrema locaux (voir section 5.3).

Un monôme de multi-degré  $\mathbf{k} = (k_1, k_2, \dots, k_m) \in (\mathbb{Z}_{\geq 0})^m$  est une fonction  $G : \mathbb{R}^m \rightarrow \mathbb{R}$  telle que pour un  $c \in \mathbb{R}$ ,  $G(\mathbf{x}) = cx_1^{k_1} x_2^{k_2} \dots x_m^{k_m} =: \mathbf{x}^{\mathbf{k}}$ . Le degré, noté  $\deg G$ , d’un monôme est  $\deg G = \sum_{i \in \overline{m}} k_i$ .

## 5.1 - Formule de Taylor

Une fonction  $F : \mathbb{R}^m \rightarrow \mathbb{R}$  est dite polynomiale ou un polynôme si elle s'écrit  $F(\mathbf{x}) = \sum_{i \in \bar{\ell}} G_i$  (où  $\ell \in \mathbb{Z}_{\geq 1}$ ). Son degré est alors  $\deg F = \max_{i \in \bar{\ell}} \deg G_i$ .

**Exemple 5.1.1.**  $4, x+y+z^2, 3x+8y^2-\pi xyz$  sont des polynômes de degré 0, 2 et 3 respectivement.  $xy+z^2/x$  n'est pas un polynôme. ♣

La preuve de la formule de Taylor-Lagrange en plusieurs variables se déduit de la version en une variable qui est ici rappelée pour le bien du lecteur.

### Théorème 5.1.2

Soit  $U \subset \mathbb{R}$  un ouvert et  $f : U \rightarrow \mathbb{R}$  une application de classe  $C^{k+1}$  sur  $U$ . Alors, pour tout  $x \in U$  et tout  $h \in \mathbb{R}$  tel que le segment  $I$  de  $x$  à  $x+h$  soit contenu dans  $U$ , il existe  $\xi \in I$  tel que

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f^{[2]}(x)h^2 + \dots + \frac{1}{k!}f^{[k]}(x)h^k + \frac{1}{(k+1)!}f^{[k+1]}(\xi)h^{k+1}.$$

En particulier, si  $R_{f,k,x}(h) := \frac{1}{(k+1)!}f^{[k+1]}(\xi)h^{k+1}$  alors  $\lim_{h \rightarrow 0} \frac{R_{f,k,x}(h)}{|h|^k} = 0$ .

Il ne s'agit plus que de faire une application de la dérivation composée pour obtenir la version en plusieurs variables.

### Théorème 5.1.3

Soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}^n$  une application de classe  $C^{k+1}$  sur  $U$ . Alors,  $\forall \mathbf{x} \in U$  et  $\forall \mathbf{h} \in \mathbb{R}^m$  tel que le segment de  $\mathbf{x}$  à  $\mathbf{x}+\mathbf{h}$  est contenu dans  $U$ ,

$$F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + (\mathbf{h} \cdot \nabla)F(\mathbf{x}) + \frac{1}{2}(\mathbf{h} \cdot \nabla)^2 F(\mathbf{x}) + \dots + \frac{1}{k!}(\mathbf{h} \cdot \nabla)^k F(\mathbf{x}) + \frac{1}{(k+1)!}(\mathbf{h} \cdot \nabla)^{k+1} F(\mathbf{x}+\tau \mathbf{h})$$

où  $\tau \in [0, 1]$ .

**DÉMONSTRATION:** Soit  $\mathbf{x}_{(0)} \in U$  le point où le développement de Taylor doit être fait, soit  $\mathbf{h} \in \mathbb{R}^m$  un vecteur (basé en  $\mathbf{x}$ ) tel que  $\forall t \in [0, 1], \mathbf{x}+t\mathbf{h} \in U$  et soit  $\mathbf{L} : [0, 1] \rightarrow \mathbb{R}^m$  l'application définie par  $\mathbf{L}(t) = \mathbf{x}+t\mathbf{h}$ . Alors la formule de Taylor-Lagrange (usuelle en une variable) de  $f(t) := F \circ \mathbf{L}(t)$  donne

$$\begin{aligned} F \circ \mathbf{L}(t) &= f(0) + f'(0)t + \frac{1}{2}f^{[2]}(0)t^2 + \dots + \frac{1}{k!}f^{[k]}(0)t^k + \frac{1}{(k+1)!}f^{[k+1]}(\tau)t^{k+1} \quad \text{où } \tau \in [0, 1], \\ &= f(0) + f'(0)t + \frac{1}{2}f^{[2]}(0)t^2 + \dots + \frac{1}{k!}f^{[k]}(0)t^k + \frac{1}{(k+1)!}f^{[k+1]}(\tau)t^{k+1}, \end{aligned}$$

en posant  $t = 1$ . Il ne reste plus qu'à vérifier que  $f^{[k]}(0) = (\mathbf{h} \cdot \nabla)^k F(\mathbf{x})$ . Ceci sera fait par induction. Pour  $k = 1$ , c'est une conséquence des dérivations composées :  $f'(0) = \mathbf{h} \cdot \nabla F(\mathbf{x}) = (\mathbf{h} \cdot \nabla)F(\mathbf{x})$  car  $\mathbf{L}_* = D_1 \mathbf{L} = \mathbf{h}$  et  $F_* = \nabla F^\top$ . Montrons que cela fonctionne aussi pour  $k \geq 2$  en supposant que  $f^{[\ell]}(\mathbf{x}) = (\mathbf{h} \cdot \nabla)^\ell F(\mathbf{x})$  pour tout  $\ell \in \overline{k-1}$ .

$$\begin{aligned} f^{[k]}(0) &= \frac{d}{dt}f^{[k-1]}(0) = \frac{d}{dt}(\mathbf{h} \cdot \nabla)^{k-1}F(\mathbf{x}+t\mathbf{h})\Big|_{t=0} \\ &= (\mathbf{h} \cdot \nabla)(\mathbf{h} \cdot \nabla)^{k-1}F(\mathbf{x}) \\ &= (\mathbf{h} \cdot \nabla)^k F(\mathbf{x}) \end{aligned} \quad \blacksquare$$



Malgré les efforts de notations pour rendre la formule ci-dessus lisible, cela ne peut pas nuire d'explicitier les quelques premiers termes dans le cas simple où  $m = 2$ ,  $F(\mathbf{x}) = F(x, y)$  et  $\mathbf{h} = \begin{pmatrix} h \\ k \end{pmatrix}$  :

$$\begin{aligned}
 F(x+h, y+k) = & F(x, y) \\
 & + hD_1F(x, y) + kD_2F(x, y) \\
 & + \frac{1}{2}(h^2D_1^2F(x, y) + 2hkD_1D_2F(x, y) + k^2D_2^2F(x, y)) \\
 & + \frac{1}{6}(h^3D_1^3F(x, y) + 3h^2kD_1^2D_2F(x, y) + 3hk^2D_1D_2^2F(x, y) + k^3D_2^3F(x, y)) \\
 & + \dots
 \end{aligned}
 \tag{5.1.4}$$

D'une ligne à la suivante, l'approximation s'améliore par l'ajout d'un polynôme homogène dont le degré est 1 de plus (*i.e.* dont tous les monômes ont le même degré).

**Définition 5.1.5.** Soit  $k \in \mathbb{Z}_{\geq 1}$  et soit  $F : U \rightarrow \mathbb{R}$  une fonction de classe  $C^k$  sur un ouvert  $U \subset \mathbb{R}^m$ . Son développement de Taylor à l'ordre  $\ell \in \overline{k}$  en  $\mathbf{x}$  est le polynôme

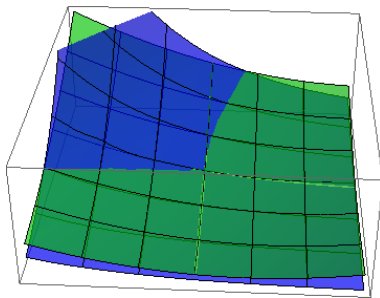
$$P_{\ell, \mathbf{x}}F(\mathbf{h}) = F(\mathbf{x}) + (\mathbf{h} \cdot \nabla)F(\mathbf{x}) + \frac{1}{2}(\mathbf{h} \cdot \nabla)^2F(\mathbf{x}) + \dots + \frac{1}{\ell!}(\mathbf{h} \cdot \nabla)^\ell F(\mathbf{x}). \quad \star$$

**Exemple 5.1.6.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction (de classe  $C^\infty$ ) définie par  $F(x, y) = e^{x^2-4y}$ . Quelques calculs rapides donnent que

$$\begin{aligned}
 D_1F(x, y) &= 2xe^{x^2-4y}, & D_2F(x, y) &= -4e^{x^2-4y}, \\
 D_1^2F(x, y) &= (2 + 4x^2)e^{x^2-4y}, & D_1D_2F(x, y) &= -8xe^{x^2-4y} \quad \text{et} \quad D_2^2F(x, y) = 16e^{x^2-4y}.
 \end{aligned}$$

Ainsi son développement de Taylor à l'ordre 2 en  $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$  est

$$P_{2, \begin{pmatrix} 2 \\ 1 \end{pmatrix}}F(h, k) = 1 + (4h - 4k) + \frac{1}{2}(18h^2 - 16hk + 16k^2),$$



(ci-contre la vraie fonction [en bleu] est comparée au développement en ce point [en vert]) tandis que le développement de Taylor à l'ordre 2 en  $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est

$$P_{2, \mathbf{0}}F(h, k) = 1 - 4k + \frac{1}{2}(2h^2 + 16k^2) \quad \clubsuit$$

Bien que ce phénomène ne soit pas nouveau, il est tout à fait possible qu'une fonction de classe  $C^\infty$  non-triviale ait un développement de Taylor trivial. L'exemple (de  $\mathbb{R} \rightarrow \mathbb{R}$ ) typique est  $f(x) = e^{-1/x^2}$ . En effet  $f^{[k]}(0) = 0$  et  $f^{[k]}$  est continue sur  $\mathbb{R}$  pour tout  $k \in \mathbb{Z}_{\geq 1}$ . En particulier, en 0, son développement de Taylor ne contient que le terme de reste. La série formée par la limite des développements est ainsi triviale (et converge automatiquement), mais n'a pour autant rien à voir avec la fonction. Comme en une variable, les applications qui peuvent s'exprimer par leur série de Taylor (sur le domaine où elles convergent) sont dites analytiques ou de classe  $C^\omega$ .

L'approximation par le développement de Taylor est, néanmoins, satisfaisante dans la plupart des cas. Soit  $F$  une fonction de classe  $C^k$  alors par la formule de Taylor-Lagrange (théorème 5.1.3), il existe  $\tau \in [0, 1]$  tel que

$$\begin{aligned} F(\mathbf{x} + \mathbf{h}) - P_{k,\mathbf{x}}F(\mathbf{h}) &= \frac{1}{k!} ((\mathbf{h} \cdot \nabla)^k F(\mathbf{x} + \tau \mathbf{h}) - (\mathbf{h} \cdot \nabla)^k F(\mathbf{x})) \\ &= \frac{\|\mathbf{h}\|^k}{k!} \left( \left( \frac{\mathbf{h}}{\|\mathbf{h}\|} \cdot \nabla \right)^k F(\mathbf{x} + \tau \mathbf{h}) - \left( \frac{\mathbf{h}}{\|\mathbf{h}\|} \cdot \nabla \right)^k F(\mathbf{x}) \right). \end{aligned}$$

Comme les dérivées d'ordre  $k$  de  $F$  sont continues, la différence ci-dessus tend rapidement vers 0 lorsque  $\mathbf{h}$  est petit. En fait, même après division par  $\|\mathbf{h}\|^k$  elle tend toujours vers 0. Ceci illustre une propriété importante de l'approximation polynomiale que permet de faire la formule de Taylor-Lagrange 5.1.3. Elle est même, en un certain sens, optimale :

### **Théorème 5.1.7**

Soit  $F : U \rightarrow \mathbb{R}$  une fonction de classe  $C^k$  sur  $U \subset \mathbb{R}^m$  et soit  $Q : \mathbb{R}^m \rightarrow \mathbb{R}$  un polynôme de degré inférieur ou égal à  $k$  tel que

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{F(\mathbf{x} + \mathbf{h}) - Q(\mathbf{h})}{\|\mathbf{h}\|^k} = 0,$$

alors  $Q$  est égal au développement de Taylor d'ordre  $k$  en  $\mathbf{x}$ .

La démonstration ne sera pas ici faite (exercice !) ; il s'agit essentiellement d'un analogue pour les dérivées d'ordres supérieurs de l'unicité de la dérivée (lorsque  $k = 1$  c'est exactement le même résultat que le théorème 3.2.5) et de l'argument esquissé ci-dessus.

Une conséquence importante, c'est que le développement de Taylor décrit (dans la mesure où il ne s'annule pas) le comportement de la fonction au voisinage du point où il se situe (cf. section 5.3).

**Exemple 5.1.8.** Une conséquence intéressante du théorème 5.1.7 est qu'il est possible de simplifier beaucoup le calcul dès qu'une approximation polynomiale est bonne. Par exemple, pour calculer le développement de  $F(x, y) = e^{x^2 - 4y}$  à l'ordre 2 en  $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$  : soit  $t = x^2 - 4y$ , alors  $t = 0$  lorsque  $x = 2$  et  $y = 1$  ; le développement de Taylor en 0 de  $e^t$  à l'ordre 2 est  $1 + t + \frac{t^2}{2}$ . En remplaçant  $t$  par  $(2+h)^2 - 4(1+k) = 4h - 4k + h^2$ , le polynôme  $1 + (4h - 4k) + \frac{1}{2}(18h^2 - 16hk + 16k^2)$  apparaît. ♣

Le théorème 5.1.7 permet aussi de percevoir une application très importante du développement de Taylor : au voisinage du point  $\mathbf{x}$ , le premier ordre où des dérivées sont non-triviales détermine le comportement dans une boule suffisamment petite autour de  $\mathbf{x}$  (e.g. la première ligne après  $F(x, y)$  dans (5.1.4) qui n'est pas identiquement nulle domine). C'est sur ce principe qu'est construite l'étude de fonction.

EXERCICE 82: *Xorqui a gagné, dans sa dernière boîte de lessive, un voyage dans un tout nouvel avion suborbital. Étant fidèle aux vraies valeurs de la tradition la plus traditionnelle, Xorqui ne considère vraie que l'heure de l'horloge grand-père de son arrière-grand-mère (un beau spécimen à pendule... l'horloge pas les ancêtres).*

*Cependant, la période d'un pendule étant donnée par  $F(L, g) = \sqrt{\frac{L}{g}}$  où  $g$  est la gravité et  $L$  la longueur du fil, il y aura une distorsion dans le comptage de la seconde car à haute altitude la*

gravité change. Xorqui veut régler son horloge pour rendre le pendule fiable et a remplacé la tige par une tige à longueur télé-réglo-programmable.

Utiliser le développement de Taylor à l'ordre 2 de  $F$  pour trouver quelle perturbation  $h$  de  $L$  peut compenser une perturbation de  $k$  de  $g$ , i.e. pour  $k$ ,  $g$  et  $L$  donnés trouver  $h$  tel que  $P_{2,(L,g)}(h,k) = 0$  (i.e. que  $F(L+h, g+k)$  est presque 0). Afin d'alléger les calculs, prenez  $g = 8$  et  $L = 2$ .

EXERCICE 83: Donner le développement de Taylor de la fonction  $F$  ci-dessous autour du point donné et à l'ordre spécifié :

- a.  $F(x,y) = \operatorname{arctg}(x+xy)$  de classe  $C^\infty$  sur  $\mathbb{R}^2$  à l'ordre 3 en  $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$  ;
- b.  $F(x,y) = \frac{\sin x}{y}$  de classe  $C^\infty$  sur  $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$  à l'ordre 3 en  $\begin{pmatrix} \pi/2 \\ 1 \end{pmatrix}$  ;
- c.  $F(x,y,z) = \frac{z}{x+y}$  de classe  $C^\infty$  sur  $\{(x,y,z) \in \mathbb{R}^3 \mid x \neq -y\}$  à l'ordre 3 en  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ .

EXERCICE 84: Montrer que les fonctions suivantes sont de classe  $C^\infty$  sur  $\mathbb{R}^3$  puis calculer les polynômes de Taylor demandé :

- a.  $F(x,y,z) = \sin(z^2 + x \cos y)$  en  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ \pi/2 \\ 0 \end{pmatrix}$  à l'ordre 3 ;
- b.  $F(x,y,z) = \operatorname{sh}(zy) \ln(1+x^2)$  en  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$  à l'ordre 6 ;
- c.  $F(x,y,z) = \operatorname{ch}(e^{xy+z^2} - 1)$  en  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$  à l'ordre 3 ;

EXERCICE 85: Calculer le polynôme de Taylor de

- a.  $F(x,y,z) = \sin(e^{xy+z^2} - 1 + x^3)$  à l'ordre 5 en  $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$  ;
- b.  $F(x,y,z) = \ln(1-y^2) \sin(x^2z + \frac{\pi}{2})$  à l'ordre 10 en  $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ .

## 5.2 Points critiques et singuliers

L'étude de fonction a pour but de comprendre le comportement de la fonction. Le plus souvent (par exemple, pour des raisons d'optimisation) les points intéressants sont les valeurs les plus grandes ou plus petites que prendra la fonction (un minimum, des minima ; un maximum, des maxima ; un extremum, des extrema). Le mot "extremum" signifie "minimum ou maximum".

**Définition 5.2.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction continue sur  $U$ . Un maximum local (resp. minimum local) de  $F$  est un point  $\mathbf{x} \in U$  tel que  $\exists r \in \mathbb{R}_{>0}$  de sorte que  $\forall \mathbf{y} \in B_r(\mathbf{x}) \cap U$ ,  $F(\mathbf{y}) \leq F(\mathbf{x})$  (resp.  $F(\mathbf{y}) \geq F(\mathbf{x})$ ).

Lorsque  $\mathbf{x}$  est un point critique de  $F$ ,  $F(\mathbf{x})$  sera appelée valeur critique de  $\mathbf{x}$ .

★



Il est impossible de trop insister sur le fait qu'un maximum ou minimum local n'est pas nécessairement un maximum ou minimum absolu. Un exercice de logique simple montrera au lecteur qu'un point n'est pas un maximum local si

$$\forall r \in \mathbb{R}_{>0} \exists \mathbf{y} \in B_r(\mathbf{x}) \cap S \text{ tel que } F(\mathbf{y}) > F(\mathbf{x});$$

*mutatis mutandis* pour ce qui est du minimum local.

Comme l'ont déjà insinué plusieurs exemples de manière plus ou moins explicite (cf. exemple 3.6.3) et comme c'était aussi le cas pour les fonctions  $\mathbb{R} \rightarrow \mathbb{R}$ , le gradient semble un indicateur (malheureusement imparfait) de la présence de minima.

### Théorème 5.2.2

Soit  $U \subset \mathbb{R}^m$  un ouvert et soit  $F : U \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . Si  $\mathbf{x}$  est un extremum de  $F$  alors  $\nabla F(\mathbf{x}) = 0$ .

Comme il est impossible de trop insister sur ce point aussi, rappelons au lecteur éventuellement étourdi qu'une implication logique  $\mathcal{A} \Rightarrow \mathcal{B}$  ne veut PAS dire que si  $\mathcal{B}$  est vraie alors  $\mathcal{A}$  est vraie. Ici en particulier, si le gradient s'annule en un point, ce point peut très bien être ni un minimum local ni un minimum absolu ni un maximum local ni un maximum absolu.

EXERCICE 86: Xorqui part toujours de sa maison tôt le matin et ne rentre jamais avant 17 heures. Soit les deux assertions suivantes :

$\mathcal{A}$  : Xorqui prend son parapluie avant de partir le matin.

$\mathcal{B}$  : Il pleut le matin.

Xorqui étant quelqu'un qui aime l'ordre, " $\mathcal{B} \Rightarrow \mathcal{A}$ " (en mots : Xorqui prend toujours son parapluie avant de partir quand il pleut le matin). Vous passez la nuit chez Xorqui et vous réveillez vers midi.

Si le parapluie n'est pas là, pouvez-vous déduire de " $\mathcal{B} \Rightarrow \mathcal{A}$ " qu'il pleuvait ce matin là ? Si le parapluie n'est pas là, pouvez-vous déduire de " $\mathcal{B} \Rightarrow \mathcal{A}$ " qu'il ne pleuvait pas ce matin là ? Justifiez et écrivez (en mots) la contraposée de  $\mathcal{B} \Rightarrow \mathcal{A}$ .

Ici la preuve sera faite en utilisant la contraposée, une autre idée de preuve se trouve aussi dans la démonstration du théorème 6.4.2.

DÉMONSTRATION PREUVE DU THÉORÈME 5.2.2: La preuve se fait simplement en montrant la contraposée : si  $\nabla F(\mathbf{x}) \neq 0$  alors  $\mathbf{x}$  n'est pas un extremum. En effet, il suffit de se rappeler de l'interprétation du gradient comme une direction de croissance (et de  $-\nabla F$  comme la direction de décroissance) pour conclure. Pour écrire cette idée rigoureusement, soit  $\mathbf{v} = \nabla F(\mathbf{x})$ , alors, comme la dérivée en  $\mathbf{x}$  existe,

$$F(\mathbf{x} + t\mathbf{v}) = F(\mathbf{x}) + t(\|\mathbf{v}\|^2 + \frac{1}{t}R_F(t\mathbf{v})) \quad \text{où} \quad \lim_{t \rightarrow 0} \frac{1}{t}R_F(t\mathbf{v}) = 0.$$

En particulier, si  $|t|$  est suffisamment petit,  $\|\mathbf{v}\|^2 + \frac{1}{t}R_F(t\mathbf{v}) > 0$  car  $\|\mathbf{v}\|^2 > 0$  ( $\mathbf{v}$  est non-nul). De plus, pour toute boule centrée en  $\mathbf{x}$  de rayon  $r$ , si  $|t| < \frac{r}{\|\mathbf{v}\|}$ , alors  $\mathbf{x} + t\mathbf{v} \in B_r(\mathbf{x})$  (car  $d(\mathbf{x}, \mathbf{x} + t\mathbf{v}) = \|t\mathbf{v}\| = |t| \|\mathbf{v}\| < r$ ). Ainsi, dans toute boule centrée en  $\mathbf{x}$  il y a un point avec une valeur légèrement supérieure (prendre  $t$  positif) et une valeur inférieure (prendre  $t$  négatif). ■

**Exemple 5.2.3.** La fonction de classe  $C^\infty$ ,  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par  $F(x, y) = (x - y)^3$  a pour gradient  $\nabla F(x, y) = \begin{pmatrix} 3(x-y)^2 \\ -3(x-y)^2 \end{pmatrix}$ . En particulier, sur toute la diagonale (les points où les deux coordonnées sont égales)  $\nabla F(x, x) = \mathbf{0}$  et  $F(x, x) = 0$ . Cependant,  $\forall x \in \mathbb{R}$ , il est possible de trouver un point proche de  $\begin{pmatrix} x \\ x \end{pmatrix}$  qui prennent une valeur plus grande ou plus petite que  $F(\mathbf{x})$ . En effet,  $F(\varepsilon + x, x) = \varepsilon^3$ , donc si  $\varepsilon < 0$  la fonction prend en  $\begin{pmatrix} \varepsilon + x \\ x \end{pmatrix}$  une valeur plus petite que  $\begin{pmatrix} x \\ x \end{pmatrix}$  et si  $\varepsilon > 0$  une valeur plus grande. ♣

**Exemple 5.2.4.** La condition de continuité des dérivées est cruciale. Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = \sqrt{x^2 + y^2} = \|(x, y)\|$ . Alors  $\nabla F(x, y) = \begin{pmatrix} x/\sqrt{x^2+y^2} \\ y/\sqrt{x^2+y^2} \end{pmatrix}$  si  $\begin{pmatrix} x \\ y \end{pmatrix} \neq \mathbf{0}$ , mais le gradient n'existe pas en  $\mathbf{0}$  (comparer les suites  $\nabla F(1/n, 0)$  et  $\nabla F(-1/n, 0)$  montre même que l'extension par continuité est impossible).  $\nabla F$  ne s'annule nulle part là où il est défini, pourtant la fonction a un minimum (absolu) en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  (cf. section 1.3 propriété No1). ♣

La terminologie suivante est courante pour l'étude des fonctions :

**Définition 5.2.5.** Soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction. Un point  $\mathbf{x} \in U$  est dit

- critique si  $F$  est dérivable en  $\mathbf{x}$  et  $\nabla F(\mathbf{x}) = \mathbf{0}$  ;
- singulier si  $F$  n'est pas dérivable en  $\mathbf{x}$  ;
- d'inflexion si  $\mathbf{x}$  est un point critique mais n'est pas un extremum local.

★

La section 5.3 donnera une méthode un peu plus raffinée qui permet parfois de conclure qu'un point admet un extremum local. Une première méthode rudimentaire est la suivante : soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction. Pour chaque  $\mathbf{x}$  point critique ou singulier, elle consiste en définir la fonction test  $G_{\mathbf{x}}(\mathbf{h}) = F(\mathbf{x} + \mathbf{h}) - F(\mathbf{x})$ . Ensuite, si pour tout  $\mathbf{h}$  assez petit  $G_{\mathbf{x}}(\mathbf{h}) \leq 0$  (resp.  $G_{\mathbf{x}}(\mathbf{h}) \geq 0$ ) alors la fonction admet un maximum local (resp. minimum local) en  $\mathbf{x}$ .

Cette méthode est évidemment difficile à réaliser : il ne s'agit en fait que d'une réécriture de la définition d'extremum local et donc, en soi, pas d'une méthode. Pour montrer que  $G_{\mathbf{x}}$  est positif ou négatif près de l'origine, seules les méthodes de manipulations classiques des inégalités sont jusqu'ici à disposition. En outre, ceci ne dit rien sur les extrema absolus.

**Exemple 5.2.6.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $2x^3 - 6xy + 3y^2$ . Tentons de lui appliquer la méthode ci-dessus.  $F$  est dérivable sur tout  $\mathbb{R}$  donc  $D = \emptyset$  et  $\nabla F(x, y) = \begin{pmatrix} 6x^2 - 6y \\ 6y - 6x \end{pmatrix}$ , donc  $C = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}\}$ .

La première fonction test  $G_0(h_1, h_2)$  n'est en fait que  $F(h_1, h_2)$ . Il faut alors remarquer que  $F(h_1, 0) = h_1^3$ , ce qui implique que dans tout voisinage de  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  il y a des points de valeurs plus grandes ( $h_1 > 0$ ) et plus petite ( $h_1 < 0$ ). Ainsi  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est un point d'inflexion.

La seconde fonction test est, en posant  $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,

$$\begin{aligned} G_{\mathbf{x}}(h_1, h_2) &= 2(1+h_1)^3 - 6(1+h_1)(1+h_2) + 3(1+h_2)^2 - (-1) \\ &= 6h_1^2 - 6h_1h_2 + 3h_2^2 + 2h_1^3 \\ &= 3(h_1 - h_2)^2 + h_1^2(3 + 2h_1). \end{aligned}$$

Pour  $h_1 > -3/2$  tous les termes sont en fait positifs (un carré étant toujours positif). Par conséquent, lorsque  $\|(h_1, h_2)\|$  est assez petit,  $G_{\mathbf{x}}(h_1, h_2) \geq 0$  et donc  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  est un minimum local.

Le moment est propice pour remarquer que  $F(x, 0) = x^3$ , ainsi, lorsque  $x \rightarrow \pm\infty$ ,  $F$  prend des valeurs arbitrairement grandes ou petites. Elle n'a donc pas d'extremum absolu. ♣

Une manière de vérifier la condition ci-dessus est de faire le développement de Taylor de  $F$  en ce point. En effet, lorsqu'il est bien défini et non-trivial, ce dernier régit le comportement de

la fonction au voisinage d'un point (voir théorème 5.1.7). Quand son expression est simple, il est possible d'obtenir une conclusion sans trop se perdre dans les détails techniques ; c'est l'objectif de la section 5.3.

## 5.3 Matrice de Hesse ou hessienne

La section précédente a permis de donner une approximation plus fine des fonctions. En particulier, lorsque  $F$  est de classe  $C^2$  et que le gradient s'annule en  $\mathbf{x}$  le théorème 5.1.7 indique que les termes qui dominent sur le comportement de la fonction sont les termes du second ordre. Ceux-ci forment un polynôme homogène de degré 2 ; un type de fonction qu'il n'est pas inespéré de pouvoir analyser.

En réalité, les polynômes homogènes de degré 2 (en  $m$  variables) portent le nom de “formes quadratiques” (sur  $\mathbb{R}^m$ ). Puisqu'il s'agit de notre motivation principale pour les étudier, la forme quadratique de  $F$  en  $\mathbf{x}$  (particulièrement si  $\mathbf{x}$  est un point critique), notée  $q_{F,\mathbf{x}}$  ou  $q_{\mathbf{x}}$ , désignera les termes de degrés 2 du polynôme de Taylor de  $F$  en  $\mathbf{x}$ .

**Exemple 5.3.1.**  $xy + \pi z^2$ ,  $2x^2 + y^2$ ,  $x_1^2 + e^8 x_2 x_3 - x_1 x_4 + x_4^2$  et  $-xy$  sont des formes quadratiques, mais  $xy + x^2 + xz^2$ ,  $xy + \frac{z^2 y}{x}$  et  $x_1 x_2 + x_4^2 \log \frac{x_3}{x_4}$  ne sont pas des polynômes homogènes de degré 2. ♣

Comme le comportement près d'un point critique est déterminé par les valeurs que peut prendre une matrice

**Définition 5.3.2.** Une forme quadratique  $q$  est dite

- Qu1. définie positive si  $\forall \mathbf{x} \in \mathbb{R}^m, q(\mathbf{x}) \geq 0$  et  $q(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = 0$  ;
- Qu2. définie négative si  $\forall \mathbf{x} \in \mathbb{R}^m, q(\mathbf{x}) \leq 0$  et  $q(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = 0$  ;
- Qu3. semi-définie positive si  $\forall \mathbf{x} \in \mathbb{R}^m, q(\mathbf{x}) \geq 0$  ;
- Qu4. semi-définie négative si  $\forall \mathbf{x} \in \mathbb{R}^m, q(\mathbf{x}) \leq 0$  ;
- Qu5. non-semi-définie si  $\exists \mathbf{x}, \mathbf{y} \in \mathbb{R}^m$  tels que  $q(\mathbf{x}) < 0$  et  $q(\mathbf{y}) > 0$ .
- Qu6. triviale si  $\forall \mathbf{x} \in \mathbb{R}^m, q(\mathbf{x}) = 0$ .



Le lecteur se convaincra sans souci que si la forme quadratique de  $F$  en un point critique  $\mathbf{x}$  est

- QE1. définie positive  $\Rightarrow \mathbf{x}$  est un minimum local ;
- QE2. définie négative  $\Rightarrow \mathbf{x}$  est un maximum local ;
- QE3. non-semi-définie  $\Rightarrow \mathbf{x}$  est un point d'inflexion ;
- QE4. semi-définie positive et non-triviale  $\Rightarrow \mathbf{x}$  n'est pas un maximum ;
- QE5. semi-définie négative et non-triviale  $\Rightarrow \mathbf{x}$  n'est pas un minimum ;

EXERCICE 87: Montrer les affirmations QE1 à 5. [Indice : Faire un développement de Taylor et utiliser le théorème 5.1.7.]

Lorsque la forme est semi-définie, il faut recourir à une étude plus poussée (voir la sous-section 5.3.i “Lorsque même la dérivée seconde ne suffit plus...”). Pour l’instant, il ne reste plus qu’à trouver/rappeler des techniques pour identifier les cas QE1-5.

Heureusement, les formes quadratiques possèdent une écriture matricielle très utile. En effet, soit  $q(\mathbf{x})$  une forme quadratique, alors

$$q(\mathbf{x}) = \sum_{i,j \in \overline{m}} Q_{ij} x_i x_j =: \mathbf{x} Q \mathbf{x}^T.$$

où  $Q_{ij} \in \mathbb{R}$ . En fait, par commutativité ( $x_i x_j = x_j x_i$ ) il est toujours possible de faire en sorte que  $Q_{ij} = Q_{ji}$ . De fait, sous cette contrainte supplémentaire, l’écriture de la forme quadratique est unique. Les réels  $Q_{ij}$  forment ainsi une matrice  $Q$  symétrique de taille  $m \times m$  (définie ci-dessus) dont le coefficient de la  $i^{\text{ème}}$  ligne et  $j^{\text{ème}}$  colonne est  $Q_{ij}$ .

**Exemple 5.3.3.** Soient

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 0 & 2 & 5 \\ 2 & 4 & 2 \\ 5 & 2 & 4 \end{pmatrix}, M_3 = \begin{pmatrix} -\pi & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix} \text{ et } M_4 = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 3 & \sqrt{2} \\ 0 & \sqrt{2} & 1 \end{pmatrix}.$$

$M_1$  est à la fois semi-définie négative et positive (la matrice triviale est toujours l’unique telle matrice).  $M_2$  est non-semi-définie (e.g. essayer  $(20, 1, 0)^T$  et  $(-20, 1, 0)^T$ ).  $M_3$  est non semi-définie (e.g. essayer  $(1, 0, 0)^T$  et  $(0, 1, 0)^T$ ).  $M_4$  est définie positive. ♣

Dans le cas particulier d’une forme quadratique provenant du polynôme de Taylor d’une fonction  $F$  (de classe  $\mathcal{C}^2$  sur un ouvert  $U \subset \mathbb{R}^m$ ) en un point critique  $\mathbf{x}$ , la matrice s’écrit

$$Q = \begin{pmatrix} D_1^2 F(\mathbf{x}) & D_1 D_2 F(\mathbf{x}) & \cdots & D_1 D_m F(\mathbf{x}) \\ D_2 D_1 F(\mathbf{x}) & D_2^2 F(\mathbf{x}) & \cdots & D_2 D_m F(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ D_m D_1 F(\mathbf{x}) & D_m D_2 F(\mathbf{x}) & \cdots & D_m^2 F(\mathbf{x}) \end{pmatrix}.$$

Elle porte le nom de matrice de Hesse ou hessienne.

Une mineure  $M'$  de taille  $k \times k$  d’une matrice  $M$  de taille  $m \times m$  est une matrice obtenue de  $M$  en supprimant  $m - k$  lignes et  $m - k$  colonnes. Une mineure principale de taille  $k \times k$  est une mineure obtenue en enlevant  $m - k$  lignes et les  $m - k$  colonnes correspondantes. Soit  $I \subset \overline{m}$  un ensemble de cardinalité  $k$ , alors  $M_I$  est la mineure principale de taille  $k \times k$  obtenue en ne gardant les lignes et colonnes dont le numéro est dans  $I$ . Si  $I = \{i_1, i_2, \dots, i_k\}$  alors

$$(M_I)_{jn} = (M)_{i_j i_n}$$

En particulier,  $M_{\overline{k}}$  est la mineure principale  $k \times k$  obtenue en ne gardant que les  $k$  premières lignes et colonnes de  $M$  :  $(M_{\overline{k}})_{jn} = (M)_{jn}, \forall j, n \in \overline{k}$ .

### Théorème 5.3.4

| Soit  $q : \mathbb{R}^m \rightarrow \mathbb{R}$  une forme quadratique, et soit  $Q$  la matrice symétrique associée. LASSE (“Les

Assertions Suivantes Sont Équivalentes”) :

DP1.  $q$  est définie positive.

DP2. Les valeurs propres de la matrice  $Q$  sont toutes  $> 0$ .

DP3. Pour tout  $k \in \overline{m}$ ,  $\text{Det } Q_{\overline{k}} > 0$ .

Le cas négatif s’obtient essentiellement en demandant que  $-Q$  soit défini positif ; voici son “LASSE” :

DN1.  $q$  est définie négative.

DN2. Les valeurs propres de la matrice  $Q$  sont toutes  $< 0$ .

DN3. Pour tout  $k \in \overline{m}$ ,  $(-1)^k \text{Det } Q_{\overline{k}} > 0$ .

Pour ce qui est du cas semi-défini, les critères sont moins commodes. Voici deux “LASSE”s :

SDP1.  $q$  est semi-définie positive.

SDP2. Les valeurs propres de la matrice  $Q$  sont toutes  $\geq 0$ .

SDP3. Pour tout sous-ensemble  $I \subset \overline{m}$ ,  $\text{Det } Q_I \geq 0$ .

SDP4. Il existe  $p \in \mathbb{Z}_{\geq 1}$  et une matrice  $R$  de taille  $m \times p$  telle que  $Q = RR^T$ .

et

SDN1.  $q$  est semi-définie négative.

SDN2. Les valeurs propres de la matrice  $Q$  sont toutes  $\leq 0$ .

SDN3. Pour tout sous-ensemble  $I \subset \overline{m}$ ,  $(-1)^{|I|} \text{Det } Q_I \geq 0$ .

SDN4. Il existe  $p \in \mathbb{Z}_{\geq 1}$  et une matrice  $R$  de taille  $m \times p$  telle que  $Q = -RR^T$ .

La démonstration de ces résultats se fait essentiellement à partir de la remarque suivante. Une matrice symétrique  $Q$  possède (toujours) des valeurs propres réelles et ses vecteurs propres peuvent (toujours) être choisis de sorte à ce qu’ils forment une base orthonormée (c’est le théorème spectral pour les matrices symétriques). Soit, pour  $i \in \overline{m}$ ,  $\lambda_i$  les valeurs propres et  $\mathbf{v}_i$  le vecteur propre associé. Alors  $\mathbf{v}_i \cdot \mathbf{v}_j = 0$  si  $i \neq j$  et  $= 1$  si  $i = j$  et  $\{\mathbf{v}_i\}_{i \in \overline{m}}$  est une base de  $\mathbb{R}^m$ . Ainsi tout  $\mathbf{x} \in \mathbb{R}^m$  s’écrit  $\mathbf{x} = \sum_{i \in \overline{m}} c_i \mathbf{v}_i$  (où  $c_i = \mathbf{x} \cdot \mathbf{v}_i$ ) et

$$\mathbf{x}Q\mathbf{x}^T = \left( \sum_{i \in \overline{m}} \lambda_i c_i \mathbf{v}_i \right) \cdot \mathbf{x} = \sum_{i \in \overline{m}} \lambda_i c_i^2.$$

Ceci suffit pour démontrer tous les critères concernant les valeurs propres. Ensuite, pour passer aux mineures associées à  $I \subset \overline{m}$ , il faut combiner ce résultat avec les matrices de projections sur les plans de coordonnées.

Enfin, il est bon de donner deux critères élémentaires :

### Proposition 5.3.5

Soit  $Q$  une matrice de taille  $m \times m$  non-triviale. Si une mineure de taille 2 de  $Q$  est non-semi-définie alors  $Q$  est non-semi-définie. Ainsi, si il existe  $\{i, j\} \subset \overline{m}$  tel que



La preuve est laissée en exercice :

EXERCICE 88: Soit  $Q$  une matrice symétrique de taille  $m \times m$ .

- Montrer que si une mineure principale de taille 2 de  $Q$  est non-semi-définie alors  $Q$  est non-semi-définie.
- Montrer qu'une matrice de taille  $2 \times 2$  est non-semi-définie et non-triviale si et seulement si son déterminant est  $< 0$

EXERCICE 89: Soit  $Q$  une matrice symétrique de taille  $m \times m$ . Supposons que pour un  $i \in \overline{m}$ ,  $Q_{ii} \neq 0$  et  $\forall j \in \overline{m} \setminus \{i\}$   $Q_{ij} = 0$ . Selon le signe de  $Q_{ii}$  et selon que la mineure principale  $Q_{\overline{m} \setminus \{i\}}$  est semi-définie positive, semi-définie négative, définie positive, définie négative ou non-semi-définie, déterminer si  $Q$  est semi-définie positive, semi-définie négative, définie positive, définie négative ou non-semi-définie.

Voici un premier exemple pour montrer que dans le cas semi-défini tout peut arriver :

**Exemple 5.3.6.** Les quatre fonctions suivantes sont de classe  $C^\infty$  sur  $\mathbb{R}^2$  et possèdent un point critique en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . L'argument qui permet de décider de quel genre de point critique il s'agit se trouve dans la sous-section 5.3.i "Lorsque même la dérivée seconde ne suffit plus..."

Soit  $F_1(x, y) = x^2 - y^2 + 2e^y - 2y$ . Le développement de Taylor à l'ordre 2 en  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est  $P_{2,0}F_1(h, k) = 2 + h^2$ , la matrice associée à la forme quadratique  $h^2$ ,  $Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ , est clairement semi-définie positive. Il s'agit en fait d'un point d'inflexion.

Soit  $F_2(x, y) = x^2 - 2xy + y^2 + x^4 + y^4$ . Le développement de Taylor à l'ordre 2 en  $\mathbf{0}$  est  $P_{2,0}F_2(h, j) = (h - k)^2$ , la matrice associée  $Q = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$  est semi-définie positive. C'est un minimum local.

Soit  $F_3(x, y) = -x^2 - 2xy - y^2 + x^4 + y^4$ . Le développement de Taylor à l'ordre 2 en  $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est aussi  $P_{2,0}F_3(h, k) = -(h + k)^2$ , la matrice associée  $Q = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$  est semi-définie négative. Cette fois-ci, c'est un point d'inflexion.

Soit  $F_4(x, y) = 2\cos x + 4\sin(xy) - 8e^y + 4y + 4\operatorname{tg} y + \frac{4}{3}y^3 - \frac{5}{72}\ln(1 + x^4)$ . Le développement de Taylor à l'ordre 2 en  $\mathbf{0}$  est  $P_{2,0}F_4(h, k) = -6 - (h - 2k)^2$ , la matrice associée  $Q = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$  est semi-définie négative. Et ce point est un maximum local.



Enfin, un exemple un peu artificiel mais non-trivial :

**Exemple 5.3.7.** Soit  $S = B_{1/2}(\mathbf{0}) \subset \mathbb{R}^4$  et  $F : S \rightarrow \mathbb{R}$  la fonction de classe  $C^\infty$  sur  $S$  définie par

$$F(x_1, x_2, x_3, x_4) = \frac{2/\sqrt{3}}{1 - x_1x_2} - 2\cos x_1 + 2\sin(x_2x_3) + 2e^{x_2^2} + x_3^2 + x_4^2 + \frac{2}{\sqrt{3}}\ln(1 + x_2x_3).$$

Son développement de Taylor à l'ordre 2 en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  indique que la forme quadratique est donnée par la matrice

$$Q = \begin{pmatrix} 1 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & 2 & 1 & 0 \\ \frac{1}{\sqrt{3}} & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Bien que légèrement longuet, il est possible de vérifier que cette matrice est définie positive. Le point  $\mathbf{0}$  est ainsi un minimum de  $F$  (il faut néanmoins vérifier que  $\nabla F(\mathbf{0}) = \mathbf{0}$ ). ♣

**Exemple 5.3.8.** Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  la fonction de classe  $C^\infty$  sur  $\mathbb{R}^3$  définie par  $F(x, y, z) = e^{-x^2-y^2} + e^{-x^2-(z-1)^2} + 2e^{xy}$ . Pour trouver les points où son gradient est nul il faut résoudre

$$\begin{aligned}\frac{\partial F}{\partial x} &= -2x(e^{-x^2-y^2} + e^{-x^2-(z-1)^2}) + 2ye^{xy} = 0 \\ \frac{\partial F}{\partial y} &= -2ye^{-x^2-y^2} + 2xe^{xy} = 0 \\ \frac{\partial F}{\partial z} &= -2(z-1)e^{-x^2-(z-1)^2} = 0\end{aligned}$$

Il y a un point critique évident en  $(0, 0, 1)^\top$ . Le calcul de son développement de Taylor à l'ordre 2 en  $\mathbf{x} = (0, 0, 1)^\top$  donne

$$P_{2,\mathbf{x}}F(h, k, \ell) = 3 - 2h^2 + 2hk - k^2 - \ell^2$$

(Il n'y a pas de termes de degré 1, comme c'est un point critique.) La matrice associée à la forme quadratique est

$$Q = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Elle est définie négative comme  $-2 < 0$ ,  $|\begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}| = 1 > 0$  et  $\text{Det } Q = -1 < 0$ ; il s'agit donc d'un maximum local de  $F$ .

L'existence d'autres points critiques est "un peu moins évidente"; elle est laissée de côté. ♣

**EXERCICE 90:** Trouver et classer (maximum/minimum local/absolu, point d'inflexion) les points critiques des fonctions suivantes :

- $F(x, y) = x \sin y$  définie sur  $\mathbb{R}^2$ ;
- $F(x_1, x_2, x_3) = x_1 x_2 x_3 - x_1^2 - x_2^2 - x_3^2$  définie sur  $\mathbb{R}^3$ .
- $F : ]\frac{\pi}{4}, \frac{7\pi}{4}[ \times ]\frac{\pi}{4}, \frac{7\pi}{4}[ \rightarrow \mathbb{R}$  définie par  $F(x, y) = \sin x \cos y$ .

**EXERCICE 91:** Dire pourquoi les fonctions suivantes sont de classe  $C^\infty$  sur  $\mathbb{R}^3$  puis trouver les points critiques :

- $F(x, y, z) = xyz e^{-(x^2+y^2+z^2)/2}$ ;
- $F(x, y, z) = x^2 + xy^2 + yz^2 - 2z$ ;
- $F(x, y, z) = \frac{xy}{2+x^4+y^4}$ ;
- $F(x, y, z) = \cos(x+z) + \sin y$ .

**EXERCICE 92:** Soit  $F : (\mathbb{R}_{>0})^2 \rightarrow \mathbb{R}$  définie par  $F(x, y) = x + 8y + \frac{1}{xy}$ . Pourquoi possède-t-elle au moins un minimum ? quel[s] est-il [sont-ils] ?

**EXERCICE 93:** Soit un prisme rectangulaire dont les côtés sont de longueur  $x, y$  et  $z \in \mathbb{R}_{>0}$ . Si le volume est fixé, quelles sont les valeurs qui minimisent la surface ? Est-ce un minimum absolu ? [Indice : utiliser la contrainte sur le volume pour réduire le nombre de variable, puis chercher un minimum à la fonction de la surface.]

EXERCICE 94: Soit  $S = \{(x, y, z)^T \in \mathbb{R}^3 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1\}$  l'ellipsoïde de demi-axes  $a$ ,  $b$ , et  $c$ . Quelle est le prisme rectangulaire de plus grand volume contenu à l'intérieur ? Est-ce un maximum absolu ? [Indice : utiliser la contrainte sur le volume pour réduire le nombre de variable, puis cherche un maximum à la fonction de volume.]

### 5.3.i Lorsque même la dérivée seconde ne suffit plus...

Lorsque la matrice est semi-définie, il y a deux situations qui peuvent survenir. La première est lorsque (miracle !) toutes les dérivées secondes s'annulent (*i.e.* la forme quadratique est triviale). Il faut alors chercher les dérivées d'ordre supérieur si la fonction est suffisamment dérivable. Si  $F$  est de classe  $C^k$  et  $k \geq 3$  est le premier ordre où les dérivées ne s'annulent pas, le comportement de la fonction est alors dominé localement par un polynôme de degré  $k$  (le développement de Taylor à l'ordre  $k$ ).

EXERCICE 95: Soit  $k > 1$  un nombre impair. Montrer que si  $F$  est une fonction de classe  $C^k$  sur un ouvert  $U \subset \mathbb{R}^m$  telle que  $\forall \ell \in \overline{k-1}, \forall i_1, i_2, \dots, i_\ell \in \overline{m}, D_{i_1} D_{i_2} \dots D_{i_\ell} F(\mathbf{x}) = 0$  et  $\exists j_1, j_2, \dots, j_k \in \overline{m}$  tel que  $D_{j_1} D_{j_2} \dots D_{j_k} F(\mathbf{x}) \neq 0$  alors  $\mathbf{x}$  est un point d'inflexion. [Indice : les polynômes homogènes de degrés impairs sont impairs.]

Si  $k > 2$  est pair, il n'y a malheureusement pas de technique aussi simple pour déterminer si un point est un extremum ou un point d'inflexion. Avec un peu de chance, il y a espoir que le polynôme sera relativement simple.

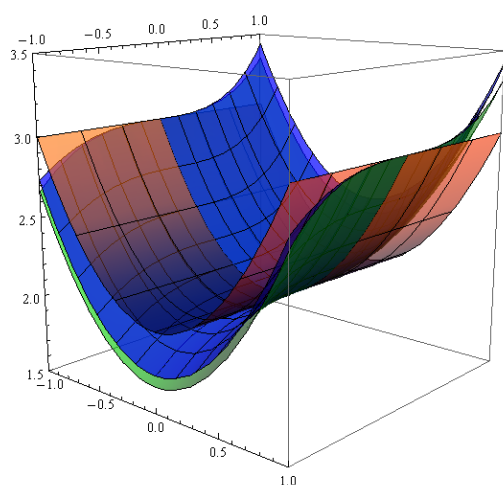
Dans l'éventualité moins miraculeuse où la forme est semi-définie mais non-triviale, il faut alors identifier les vecteurs  $\mathbf{w} \in \mathbb{R}^m$  tels que  $q(\mathbf{w}) = 0$  (par homogénéité il suffit d'en prendre un par direction) ; soit  $W$  l'ensemble de ces vecteurs. Il faut alors calculer des ordres de dérivation supplémentaires, *c'ad.* calculer  $p_k$  le polynôme homogène de degré  $k$  formé par les termes de degrés  $k$  dans le développement de Taylor. Ensuite, il faut tenter de voir si  $p_k$  est (strictement) positif, (strictement) négatif ou peut prendre des valeurs (str.) positive ET (str.) négative sur  $W \setminus \{0\}$  (par homogénéité, il est suffisant de regarder les vecteurs de norme 1 dans  $W$ ).

En bref, voici une méthode. Soit  $\varepsilon = \pm 1$  le signe de la forme quadratique semi-définie.

1. Choisir un  $\mathbf{w}$  dans  $W$ .
2. Trouver le plus petit  $k$  tel que  $p_k(\mathbf{w}) \neq 0$ . [Attention : il se peut que ce  $k$  n'existe pas, *c'ad.* il existe des fonctions  $C^\infty$  qui ne sont pas  $C^\omega$ .]
3. Si  $k_{\mathbf{w}}$  est impair,  $\mathbf{w}$  est un point d'inflexion (*cf.* exercice 95).
4. Si  $k_{\mathbf{w}}$  est pair et  $p_k(\mathbf{x})$  a un signe qui n'est pas le même que  $\varepsilon$  alors  $\mathbf{w}$  est un point d'inflexion.
5. Sinon passer à la prochaine direction présente dans  $W$ . Si  $\forall \mathbf{w}$ ,  $k_{\mathbf{w}}$  a pu être trouvé et les clauses 3 et 4 ne s'appliquent pas, c'est que le point est un extremum local (maximum si  $\varepsilon = -1$ , minimum si  $\varepsilon = +1$ ).
6. Si  $\exists \mathbf{w} \in W$  tel que  $k_{\mathbf{w}}$  n'existe pas (le cas des fonctions infiniment plates), alors la conclusion ne peut être précisée (*i.e.* c'est soit un point d'inflexion soit un maximum/minimum local selon que  $\varepsilon = \mp 1$ ).

Reprenons les fonction de l'exemple 5.3.6

**Exemple 5.3.9.** Pour  $F_1$ , la forme quadratique s'annule sur la droite  $h = 0$ . Le polynôme homogène de degré 3 qu'il faut ajouter pour obtenir le développement à l'ordre 3 est  $\frac{1}{3}k^3$ . Ainsi,  $F(0, k) - F(0, 0)$  sera positif pour  $k$  positif, mais négatif pour  $k$  négatif, il s'agit donc d'un point d'inflexion. Sur le dessin un peu plus bas sont représentés, en rouge l'approximation de  $F_1$  par un polynôme de degré 2 (qui ne montre pas la nature du point critique), en vert celle par un polynôme de degré 3 (qui montre que c'est un point d'inflexion), puis en bleu le graphe de la fonction (qui montre effectivement que ce n'est pas un point critique, mais bien un point d'inflexion).

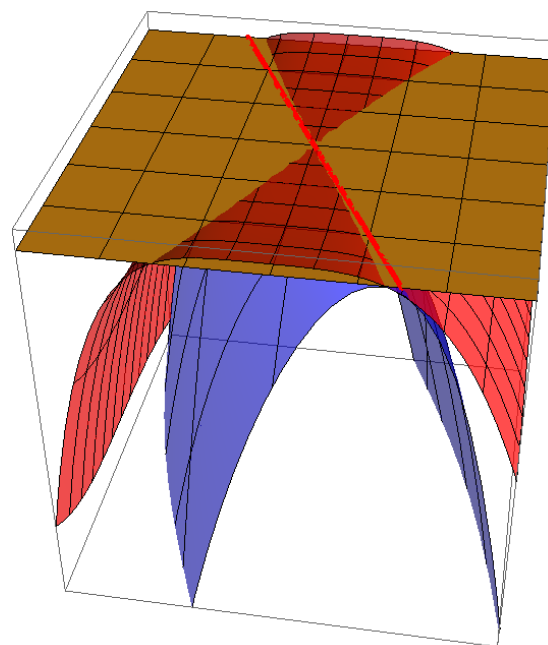


Pour  $F_2$ , la forme quadratique s'annule lorsque  $h = k$ . Le développement de Taylor à l'ordre 3 est identique à celui de l'ordre 2 ; il faut ajouter le polynôme (homogène de degré 4)  $h^4 + k^4$  pour avoir la prochaine strate non-triviale du développement de Taylor. Or pour  $h = k$  celle-ci est toujours positive, ainsi le point est un maximum local.

Dans le cas de  $F_3$ , il faut aussi aller à l'ordre 4 pour avoir la strate non-triviale,  $h^4 + k^4$  mais cette fois-ci elle est positive alors que la matrice est semi-définie négative ; il s'agit d'un point

d'inflexion.

Finalement,  $F_4$  a aussi une strate de degré 3 triviale. Celle de degré 4 est  $\frac{1}{72}h^4 - \frac{1}{3}k^4$ . La direction où la forme quadratique s'annule était donnée par  $h = 2k$  ; en remplaçant cette valeur, la strate de degré 4 sur cette droite vaut  $\frac{2}{9}k^4 - \frac{1}{3}k^4 = -\frac{1}{9}k^4$ . C'est-à-dire des valeurs négatives. Ainsi, le point est un maximum.



La représentation graphique est ici plus délicate. Ci-dessus : en bleu le graphe de la fonction  $F_4$ , en rouge la fonction moins son approximation d'ordre 2, en brun le plan de la valeur 0 (*i.e.*  $z = 0$ ). De fait, même si  $G = F_4 - P_{2,0}$  est parfois au dessus de 0,  $G$  ne prend que des valeurs négatives le long de la droite qui n'est pas dominée par la partie quadratique d'ordre 2 (en rouge vif). Ainsi il s'agit bien d'un maximum. ♣

## Chapitre 6

# Résolution et étude II

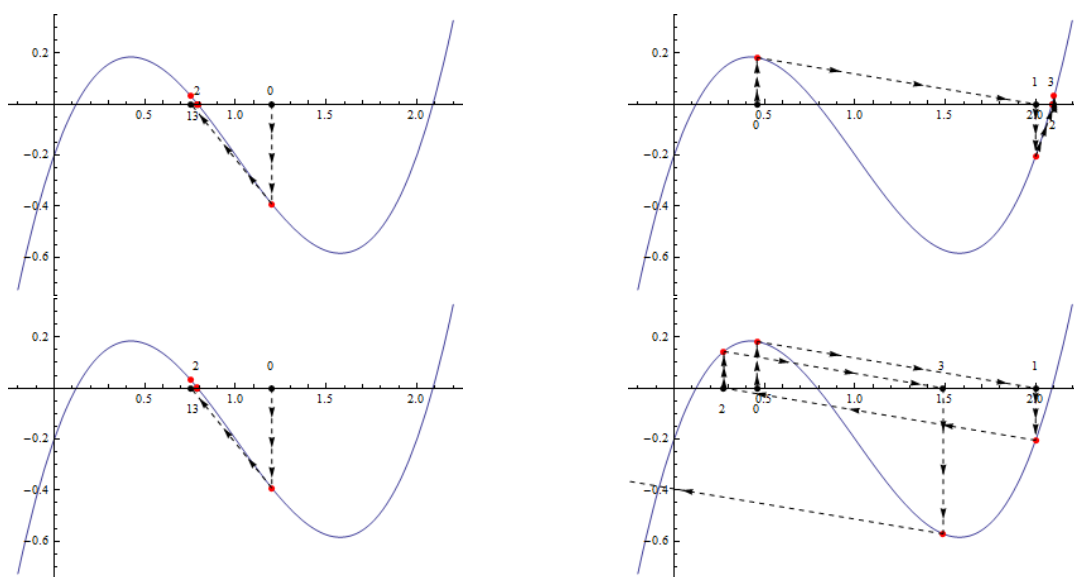
*“On ne parle jamais assez de variétés au théâtre.”*

-UN CRITIQUE SURPRIT SUR CE POINT

Ce chapitre a pour but de faire une démonstration correcte des multiplicateurs de Lagrange (c'ad. de l'étude de fonction sous avec une ou plusieurs contraintes, typiquement non-linéaires). Pour y parvenir, il d'abord établir une méthode pour la résolution des équations non-linéaires, *i.e.* du type  $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ . Il s'agit essentiellement de la version à plusieurs variables de la méthode de Newton. De là, le théorème d'inversion locale sera démontrée (avec quelques boîtes noires) et le théorème des fonctions implicites en sera déduit. Ces deux théorèmes sont le fondement de la géométrie différentielle et de la théorie des variétés. Pour nos ambitions dans ce cours, ils sont nécessaires pour la démonstration de la méthode de Lagrange, et, plus tard, pour l'étude des surfaces paramétrées.

### 6.1 Méthode de Newton

La méthode de Newton est une méthode élémentaire (et pourtant efficace) pour former une suite qui convergera vers une solution d'une équation relativement régulière. La procédure en plusieurs variables est très similaire à la procédure en une variable qui est ici brièvement rappelée. Soit  $U \subset \mathbb{R}$  un ouvert et soit  $f : U \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . Le problème est de trouver une solution à l'équation  $f(x) = 0$ . Soit  $x_0$  tel que  $f(x_0)$  est petit et  $f'(x_0) \neq 0$ . Un dessin où une solution à l'équation existe proche de  $x_0$  convaincra le chaland que si  $x_1$  est donné par  $f(x_0) = (x_0 - x_1)f'(x_0)$ , ou autrement écrit  $x_1 = x_0 - f'(x_0)^{-1}f(x_0)$ , alors  $x_1$  est plus proche de la solution. L'idée est alors d'itérer :  $x_{n+1} = x_n - f'(x_n)^{-1}f(x_n)$  converge raisonnablement vers une solution. Entre parenthèses, même dans le cas à une variable, la convergence n'est pas automatique pour tout point de départ. Voici quelques exemples pour une simple fonction polynomiale de degré 3 :



Les deux dessins du haut représentent l'algorithme pour la méthode de Newton (usuelle) et ceux d'en bas pour la méthode de Newton modifiée (avec les mêmes points de départ). Le domaine de convergence pour la méthode de Newton est en général bien meilleur (*cf.* les deux dessins de droite), mais pour les bons cas (essentiellement, lorsque le point de départ est près d'une solution et que la dérivée ne varie pas trop) la différence est faible (*cf.* les graphiques de gauche ci-dessus).

Parfois, pour sauver du temps de calcul, l'itération utilisée est  $x_{n+1} = x_n - f'(x_0)^{-1} f(x_n)$  ; cette itération, dite méthode de Newton modifiée, est particulièrement courue en grande dimension (d'autant plus lorsqu'elle est infinie). En plusieurs variables, la méthode est essentiellement identique, mais il faut remplacer  $f'(x)$  par la matrice de Jacobi. Celle-ci n'étant la plupart du temps pas une matrice carrée, et son inverse étant parfois longue à établir, il sera plus commode d'utiliser la version dite modifiée de la méthode de Newton.

La problématique est la suivante : soit  $n \geq m \in \mathbb{Z}_{\geq 1}$ , soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application de classe  $\mathcal{C}^2$  sur  $U \subset \mathbb{R}^m$ , et soit  $\mathbf{y} \in \mathbb{R}^n$ , comment trouver une solution  $\mathbf{x}$  à  $\mathbf{F}(\mathbf{x}) = \mathbf{y}$  ?

L'équation n'a évidemment pas une solution pour n'importe quel  $\mathbf{y}$ , l'hypothèse de base ici est qu'elle aura au moins une presque solution (puis quelques hypothèses techniques sur son comportement près de la presque solution).

La méthode en plusieurs variables est essentiellement la même qu'en une seule variable. Ce qui motive une petite parenthèse sur l'inversion de matrice non-carrée : dans notre cas, la dérivée est représentée par une matrice de taille  $n \times m$  (la jacobienne est de taille  $n \times m$ , *cf.* théorème 3.3.3). Pour une matrice  $M$  de taille  $n \times m$ , il faut distinguer entre un inverse à gauche ( $G_M M = \text{Id}_{\mathbb{R}^m}$ , qui n'existe pas quand  $n > m$ ) et l'inverse à droite ( $M D_M = \text{Id}_{\mathbb{R}^n}$ , qui n'existe pas si  $m > n$ ). Une possibilité pour trouver un inverse à gauche est la matrice

$$G_M = (M^T M)^{-1} M^T,$$

dans la mesure où  $M^T M$  est inversible. Il est aussi possible, par exemple, de chercher un mineur de la matrice  $M$  de déterminant non-nul, de l'inverser, et de construire un  $G_M$  à partir de là.

L'itération de Newton s'écrit de manière plus commode comme suit : soit  $\mathbf{x}$  le point de départ de l'itération et, pour  $i \in \mathbb{Z}_{\geq 0}$ , les vecteurs  $\mathbf{h}_{(i)} \in \mathbb{R}^m$  (basés en  $\mathbf{x}$ ) avec  $\mathbf{h}_{(0)} = \mathbf{0}$ . Alors l'itération s'exprime (en remplaçant  $x_n$  par  $\mathbf{x} + \mathbf{h}_{(n)}$ )

$$\mathbf{h}_{(n+1)} = \mathbf{h}_{(n)} - G_{F_*(\mathbf{x})} \mathbf{F}(\mathbf{x} + \mathbf{h}_{(n)}) =: \mathbf{N}(\mathbf{h}_{(n)}),$$

où  $G_{F_*(\mathbf{x})}$  est l'inverse à gauche de la matrice de Jacobi  $F_*(\mathbf{x})$ . Le résultat principal de cette section est un théorème donnant une condition qui assure la convergence de cette itération. Il est nécessaire de faire quelques préparatifs. Le premier est un joli lemme très utile et polyvalent : un théorème du point fixe pour les applications contractantes.

**Définition 6.1.1.** Soit  $(X, d)$  un espace muni d'une distance. Une application  $f : X \rightarrow X$  est dite contractante si  $\exists \lambda \in ]0, 1[$  tel que  $\forall x, x' \in X, d(f(x), f(x')) \leq \lambda d(x, x')$ . ★

**Lemme 6.1.2 (Théorème du point fixe de Banach)**

Soit  $(X, d)$  un espace métrique complet. Toute application  $f : X \rightarrow X$  continue et contractante possède un unique point fixe, c'ad.  $\exists x \in X$  tel que  $f(x) = x$ . De plus, ce point fixe est unique et,  $\forall x_0 \in X, d(x_0, x) \leq \frac{1}{1-\lambda} d(x_0, f(x_0))$ .

**DÉMONSTRATION :** Soit  $\lambda$  le réel de la définition 6.1.1. Soit  $x_0 \in X$  un point quelconque, et soit  $\{x_n\}_{n \in \mathbb{Z}_{\geq 0}}$  la suite définie par  $x_{n+1} = f(x_n)$ . Soit  $k, \ell \in \mathbb{Z}_{\geq 1}$ , alors  $d(x_k, x_\ell) = d(f(x_{k-1}), f(x_{\ell-1})) \leq \lambda d(x_{k-1}, x_{\ell-1})$ . L'inégalité peut être itérée pour obtenir, si  $k > \ell$ , que  $d(x_k, x_\ell) \leq \lambda^\ell d(x_{k-\ell}, x_0)$ . L'inégalité du triangle donne ensuite que

$$d(x_n, x_0) \leq \sum_{i=0}^{n-1} d(x_{i+1}, x_i) \leq \sum_{i=0}^{n-1} \lambda^i d(x_1, x_0) = \frac{1 - \lambda^n}{1 - \lambda} d(x_1, x_0).$$

Par conséquent, si  $k \in \mathbb{Z}_{>\ell}$ , alors

$$d(x_k, x_\ell) \leq \lambda^\ell d(x_{k-\ell}, x_0) \leq \frac{\lambda^\ell (1 - \lambda^{k-\ell})}{1 - \lambda} d(x_1, x_0) \leq \lambda^\ell \frac{d(x_1, x_0)}{1 - \lambda}.$$

De là, il apparaît que c'est une suite de Cauchy (i.e.  $\forall \varepsilon > 0, \exists N_\varepsilon$  tel que,  $\forall k, \ell > N_\varepsilon$ , on ait  $d(x_k, x_\ell) < \varepsilon$ ). Pour trouver  $N_\varepsilon$ , il suffit de s'assurer que

$$\lambda^{N_\varepsilon} \frac{d(x_1, x_0)}{1 - \lambda} < \varepsilon \Leftrightarrow N_\varepsilon > \frac{\ln \varepsilon + \ln(1 - \lambda) - \ln d(x_1, x_0)}{\ln \lambda}.$$

Ainsi la suite converge (car  $X$  est complet), et le point vers lequel elle converge a la propriété suivante :

$$\lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} f(x_k) = f\left(\lim_{k \rightarrow \infty} x_k\right).$$

Autrement dit, la limite est en fait un point fixe de  $f$ , ce qui mène à la conclusion de l'existence.

Si  $f$  avait deux points fixes, disons  $\bar{x}$  et  $\tilde{x}$  alors  $d(\bar{x}, \tilde{x}) = d(f(\bar{x}), f(\tilde{x})) \leq \lambda d(\bar{x}, \tilde{x})$  une contradiction comme  $\lambda < 1$ . Ainsi, l'unicité est établie. Notons  $x_\infty$  ce point limite ;  $x_\infty = \lim_{n \rightarrow \infty} x_n$ .

Finalement, en évaluant la série dont les termes sont les distances successives  $d(x_{n+1}, x_n)$ , l'estimation ci-dessus sur les distances permet de conclure que  $d(x_0, x_\infty) \leq \frac{1}{1-\lambda} d(x_0, f(x_0))$  et  $\forall n \in \mathbb{Z}_{\geq 1}, d(x_0, x_n) \leq \frac{1}{1-\lambda} d(x_0, f(x_0))$ . ■

Un autre ingrédient de la preuve de la convergence de la méthode de Newton est une version du théorème des accroissements finis pour les fonctions à valeurs vectorielles :

**Lemme 6.1.3 (Inégalité des accroissements finis)**

Soit  $U$  un ouvert de  $\mathbb{R}^m$ , soit  $\mathbf{x}, \mathbf{y} \in U$ , soit  $\gamma : [0, 1] \rightarrow U$  la courbe paramétrée (de classe  $C^1$ ) définie par  $\gamma(t) = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$  et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  tel que  $\mathbf{F}$  est dérivable en  $\gamma(t) = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$  pour tout  $t \in [0, 1]$ . Si de plus  $\forall t \in [0, 1], \exists K \in \mathbb{R}_{>0}$  tel que  $\|\mathbf{F}_*(\gamma(t))\|_{\rightarrow} \leq K$  alors  $\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\| \leq K\|\mathbf{x} - \mathbf{y}\|$ .

DÉMONSTRATION: Soit  $\varepsilon \in \mathbb{R}_{>0}$  et soit  $S = \{s \in [0, 1] \mid \|\mathbf{F}(\gamma(s)) - \mathbf{F}(\mathbf{x})\| \leq (K + \varepsilon)s\|\mathbf{x} - \mathbf{y}\|\}$ . Par continuité de  $\mathbf{F}$ , si  $t$  est le supremum de  $S$ , alors  $t \in S$ . Supposons que  $t < 1$ , alors il existe une suite de réels  $r_k$  tendant vers 0 tels que

$$\begin{aligned} (K + \varepsilon)(t + r_k)\|\mathbf{x} - \mathbf{y}\| &< \|\mathbf{F}(\gamma(t + r_k)) - \mathbf{F}(\mathbf{x})\| \\ &\leq \|\mathbf{F}(\mathbf{x} + (t + r_k)(\mathbf{y} - \mathbf{x})) - \mathbf{F}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\| + \|\mathbf{F}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - \mathbf{F}(\mathbf{x})\| \\ &\leq \|\mathbf{F}(\mathbf{x} + (t + r_k)(\mathbf{y} - \mathbf{x})) - \mathbf{F}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\| + (K + \varepsilon)t\|\mathbf{x} - \mathbf{y}\| \end{aligned}$$

D'où  $(K + \varepsilon)r_k\|\mathbf{x} - \mathbf{y}\| < \|\mathbf{F}(\mathbf{x} + (t + r_k)(\mathbf{y} - \mathbf{x})) - \mathbf{F}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\|$ . En divisant par  $r_k$  des deux côtés de l'inégalité, et en passant à la limite :

$$\begin{aligned} (K + \varepsilon)\|\mathbf{x} - \mathbf{y}\| &\leq \|\mathbf{F}_*(\gamma(t))(\mathbf{x} - \mathbf{y})\| \\ &\leq K\|\mathbf{x} - \mathbf{y}\|. \end{aligned}$$

Il s'agit là d'une contradiction. Par conséquent,  $t = 1$ . Puisque ceci est vrai pour tout  $\varepsilon > 0$ , la conclusion s'ensuit. ■

Un lecteur assidu pourra essayer de démontrer ce résultat en remplaçant  $\gamma$  par une courbe arbitraire de  $\mathbf{x}$  à  $\mathbf{y}$  et  $\|\mathbf{x} - \mathbf{y}\|$  par la longueur de  $\gamma$ .

L'idée est maintenant de montrer que si la dérivée de l'application de Newton est assez petite, elle sera contractante (lemme 6.1.3) et  $\mathbf{N}$  aura alors un point fixe (lemme 6.1.2). Si  $\mathbf{N}(\mathbf{h}) = \mathbf{h}$ , il est facile de voir que  $\mathbf{F}(\mathbf{x} + \mathbf{h}) = \mathbf{0}$ .

**Théorème 6.1.4**

Soit  $S \subset \mathbb{R}^m$ , soit  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  et soit  $\mathbf{x}$  telle que  $\mathbf{F}_*(\mathbf{x})$  est inversible à gauche. Soit  $G := G_{\mathbf{F}_*(\mathbf{x})}$  l'inverse à gauche de  $\mathbf{F}_*(\mathbf{x})$ , soit  $K = \|G\|_{\rightarrow} \in \mathbb{R}_{>0}$ , soit  $\varepsilon = \|\mathbf{F}(\mathbf{x})\| \in \mathbb{R}_{>0}$ . Supposons que  $\exists r, \delta \in \mathbb{R}_{>0}$  tels que  $\forall \mathbf{y} \in B_r(\mathbf{x}), \|\mathbf{F}_*(\mathbf{y}) - \mathbf{F}_*(\mathbf{x})\|_{\rightarrow} \leq \delta$ . Si  $0 < \frac{\varepsilon}{K-1-\delta} < r$ , alors la suite définie par l'itération de Newton  $\mathbf{h}_{(n+1)} = \mathbf{N}(\mathbf{h}_{(n)})$  converge vers une unique solution dans un voisinage de  $\mathbf{x} : \mathbf{F}(\mathbf{x} + \mathbf{h}_{(\infty)}) = \mathbf{0}$ .

DÉMONSTRATION: Soit  $\mathbf{N}(\mathbf{h}_{(n)}) = \mathbf{h}_{(n+1)} = \mathbf{h}_{(n)} - G\mathbf{F}(\mathbf{x} + \mathbf{h}_{(n)})$  comme plus haut. Alors la dérivée de  $\mathbf{N}$  est

$$\mathbf{N}_*(\mathbf{h}) = \text{Id} - G\mathbf{F}_*(\mathbf{x} + \mathbf{h}) = G(\mathbf{F}_*(\mathbf{x}) - \mathbf{F}_*(\mathbf{x} + \mathbf{h}))$$

pour tout  $\mathbf{h}$  tel que  $\|\mathbf{h}\| < r$ . Ainsi

$$\|\mathbf{N}_*(\mathbf{h})\| \leq \|G\|_{\rightarrow} \|\mathbf{F}_*(\mathbf{x}) - \mathbf{F}_*(\mathbf{x} + \mathbf{h})\|_{\rightarrow} \leq K\delta.$$



Par hypothèse  $K^{-1} - \delta > 0$ , d'où  $\delta K < 1$ . Le lemme 6.1.3 donne que,  $\forall \mathbf{h}, \mathbf{h}' \in B_r(\mathbf{0})$ ,  $\|\mathbf{N}(\mathbf{h}) - \mathbf{N}(\mathbf{h}')\| \leq \delta K \|\mathbf{h} - \mathbf{h}'\|$ . L'hypothèse implique aussi que  $K\varepsilon < r$ , ainsi  $\mathbf{h}_{(0)} = \mathbf{0}, \mathbf{h}_{(1)} = \mathbf{N}(\mathbf{0}) \in B_r(\mathbf{0})$  (car  $\|\mathbf{N}(\mathbf{0})\| = \|\mathbf{G}\mathbf{F}(\mathbf{x})\| \leq \|G\| \|\mathbf{F}(\mathbf{x})\| \leq K\varepsilon$ ).

Ceci est important car il faut aussi vérifier que la suite reste à l'intérieur de la boule où l'application est contractante. Or la preuve du lemme 6.1.2 donne aussi que les membres suivants de la suite resteront à distance  $\frac{K\varepsilon}{1-K\delta} < r$  de  $\mathbf{0}$ . Ainsi, la suite commençant en  $\mathbf{0}$  convergera, et par continuité il existe un voisinage de  $\mathbf{0}$  autour duquel toute suite convergera vers le même point fixe. ■



Bien qu'elle puisse sembler anodine, la méthode de Newton est l'ingrédient (ou l'idée) crucial(e) derrière deux théorèmes équivalents qui servent de fondement à la géométrie différentielle et à la démonstration des multiplicateurs de Lagrange. C'est le sujet des prochaines sections.

## 6.2 Théorème des applications implicites

La question de la dimension, d'apparence triviale, est beaucoup plus subtile qu'il y paraît à première vue ; la section 6.A présente quelque peu le sujet. D'un point de vue dimensionnel, le bon type d'objet à considérer sont les homéomorphismes. Pour préserver plus de structures (comme il est, entre autres, nécessaire en géométrie différentielle) il faut parler de difféomorphisme.

**Définition 6.2.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert et  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  une application. Soit  $k \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$ .  $\mathbf{F}$  est un difféomorphisme  $C^k$  si  $\mathbf{F}$  est une bijection de  $U$  dans  $\mathbf{F}(U)$ ,  $\mathbf{F}$  est de classe  $C^k$  sur  $U$  et l'inverse de  $\mathbf{F}$  est de classe  $C^k$  sur  $\mathbf{F}(U)$ . En particulier, lorsque  $k = 0$ ,  $\mathbf{F}$  est alors appelée un homéomorphisme (rappel :  $C^0$  sur  $U$  = continue sur  $U$ ).

Une application  $\mathbf{F}$  est dite un difféomorphisme  $C^k$  local en  $\mathbf{a}$  si c'est un difféomorphisme  $C^k$  sur un voisinage de  $\mathbf{a}$ . ★

Si  $\mathbf{F} : U \rightarrow \mathbf{F}(U)$  est un difféomorphisme  $C^1$  et si  $\mathbf{G} : \mathbf{F}(U) \rightarrow U$  est sa réciproque alors  $\mathbf{G} \circ \mathbf{F} = \text{Id}$ . Par la règle de dérivation composée,  $\mathbf{G}_*(\mathbf{F}(\mathbf{x}))\mathbf{F}_*(\mathbf{x}) = \text{Id}$  ou autrement dit  $\mathbf{G}_*(\mathbf{F}(\mathbf{x})) = \mathbf{F}_*(\mathbf{x})^{-1}$  (l'inverse au sens matriciel). Ainsi, un difféomorphisme  $C^1$  possède toujours une jacobienne inversible. Le théorème d'inversion locale exprime que ce critère est, dans la mesure du possible, déterminant.

### Théorème 6.2.2 (Théorème d'inversion locale)

Soit  $U$  un ouvert de  $\mathbb{R}^m$  et  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  une application de classe  $C^1$ . Si  $\exists \mathbf{x} \in U$  tel que  $\mathbf{F}_*(\mathbf{x})$  est inversible alors  $\mathbf{F}$  est un difféomorphisme  $C^1$  local en  $\mathbf{x}$ .

*Parenthèse :* Ce théorème est essentiellement la version non-linéaire du résultat (bien connu) sur les matrices (les applications linéaires) disant qu'une matrice est inversible si et seulement si son déterminant est non-nul.

**DÉMONSTRATION DÉMONSTRATION PARTIELLE ::** Il s'agit essentiellement d'une conséquence de la convergence de la méthode de Newton. Pour  $\mathbf{y}$  proche de  $\mathbf{F}(\mathbf{x})$ , il faut montrer qu'il y a une solution unique à  $\mathbf{F}(\mathbf{z}) = \mathbf{y}$  où  $\mathbf{z}$  est dans un voisinage de  $\mathbf{x}$ . Ceci est non-trivial car même si  $\mathbf{F}^{-1}$  est toujours définie en tant qu'application image réciproque, l'image réciproque d'un point peut être

plusieurs valeurs. Pour se remettre dans le cadre de la méthode de Newton, il suffit d'introduire la fonction secondaire  $\mathbf{G}(\mathbf{z}) = \mathbf{F}(\mathbf{z}) - \mathbf{y}$ . Ainsi, lorsque  $\mathbf{G}(\mathbf{z}) = 0$ ,  $\mathbf{z}$  satisfait  $\mathbf{F}(\mathbf{z}) = \mathbf{y}$ .

Soit  $K = \|\mathbf{G}_*(\mathbf{x})^{-1}\|_{\rightarrow}$ . Comme  $\mathbf{G}_*$  est continue,  $\forall \delta > 0 \exists r \in \mathbb{R}_{>0}$  tel que  $\forall \mathbf{z} \in B_r(\mathbf{x}), \|\mathbf{G}_*(\mathbf{z}) - \mathbf{G}_*(\mathbf{x})\|_{\rightarrow} \leq \delta$ . Prendre  $\delta = \frac{1}{2K}$  fera l'affaire. Maintenant, si  $\varepsilon = \frac{r}{2K}$  et  $\mathbf{y} \in B_\varepsilon(\mathbf{F}(\mathbf{x}))$ , le théorème 6.1.4 dit que l'itération commençant en  $\mathbf{x}$  convergera : il y a donc une solution à  $\mathbf{G}(\mathbf{z}) = 0$  si  $\mathbf{y} \in B_\varepsilon(\mathbf{F}(\mathbf{x}))$ , c'est-à-dire un inverse à  $\mathbf{F}$  sur  $B_r(\mathbf{x}) \cap \mathbf{F}^{-1}(B_\varepsilon(\mathbf{F}(\mathbf{x})))$ .

Pour compléter la démonstration, il faudrait montrer que la dépendance de  $\mathbf{z}$  en  $\mathbf{y}$  est  $\mathcal{C}^1$ . Il faut en fait raffiner la preuve du théorème 6.1.4 en se servant du fait que  $\mathbf{F}_*$  est continue. Pour les détails, voir le livre de C. H. Edwards [Jr.], "Advanced calculus of several variables", cote G99 dans la bibliothèque de l'institut de mathématiques de l'Unine. ■

Le premier exemple montre que les hypothèses sont bien minimales.

**Exemple 6.2.3.** La fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  donnée par  $f(x) = x^3$  est bijective et de classe  $\mathcal{C}^1$  (même  $\mathcal{C}^\infty$ ) mais sa réciproque  $f^{-1}(y) = y^{1/3}$  n'est pas dérivable en  $y = 0$ . De fait,  $f'(0) = 0$ . ♣

**Exemple 6.2.4.** Soit  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  définie par  $\mathbf{F}(x, y) = \begin{pmatrix} xe^y \\ y + \cos x \end{pmatrix}$ , alors

$$\mathbf{F}_*(x, y) = \begin{pmatrix} e^y & xe^y \\ -\sin x & 1 \end{pmatrix} \Rightarrow \text{Det } \mathbf{F}_*(x, y) = e^y(1 + x \sin x).$$

En particulier  $\mathbf{F}$  est un difféomorphisme  $\mathcal{C}^1$  local en  $\begin{pmatrix} \pi \\ 0 \end{pmatrix}$ , et la jacobienne de son inverse  $\mathbf{G}$  en ce point est  $\mathbf{G}_*(\pi, -1) = \mathbf{F}_*(0, \pi)^{-1} = \begin{pmatrix} 1 & -\pi \\ 0 & 1 \end{pmatrix}$ . ♣

EXERCICE 96: Dire si les applications suivantes sont inversibles sur un voisinage du point précisé, et, si oui, donner la dérivée.

- $\mathbf{F}(x, y) = \begin{pmatrix} e^y \cos x \\ \ln(x+y) \end{pmatrix}$  au voisinage de  $\begin{pmatrix} \pi/2 \\ 0 \end{pmatrix}$ ;
- $\mathbf{F}(x, y, z) = \begin{pmatrix} \sin x + xy \\ \arctg(xe^y) \\ z^3 + y^3 + \ln(1+x^2) \end{pmatrix}$  en  $\begin{pmatrix} \pi/2 \\ 1 \\ 1 \end{pmatrix}$ ;
- $\mathbf{F}(x, y, z) = \begin{pmatrix} x^2 + 2xy \\ y^3 + xyz \\ 2yz + z^2 \end{pmatrix}$  en  $\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$ .

EXERCICE 97: Soit  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  l'application déterminée par  $\mathbf{F}(x, y) = \begin{pmatrix} y^2 + x(1-x^2) \\ xy \end{pmatrix}$ . Quel est le plus grand ouvert  $U \subset \mathbb{R}^2$  tel que  $\forall \begin{pmatrix} x \\ y \end{pmatrix} \in U$ ,  $\mathbf{F}$  est un difféomorphisme  $\mathcal{C}^1$  local en  $\begin{pmatrix} x \\ y \end{pmatrix}$ .  $\mathbf{F}$  est-elle alors un difféomorphisme  $\mathcal{C}^1$  de  $U$  vers  $\mathbf{F}(U)$  ?

EXERCICE 98: Soit  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  l'application déterminée par  $\mathbf{F}(x, y) = \begin{pmatrix} e^{-y} \cos x \\ e^{-y} \sin x \end{pmatrix}$ .

- Calculer la matrice de Jacobi en  $\begin{pmatrix} x \\ y \end{pmatrix}$  et donner son déterminant.
- Quel est le plus grand ouvert  $U \subset \mathbb{R}^2$  tel que  $\forall \begin{pmatrix} x \\ y \end{pmatrix} \in U$ ,  $\mathbf{F}$  est un difféomorphisme  $\mathcal{C}^1$  local en  $\begin{pmatrix} x \\ y \end{pmatrix}$  ?
- Montrer que  $\mathbf{F}$  n'est pas un difféomorphisme  $\mathcal{C}^1$  de  $U$  vers  $\mathbf{F}(U)$ . [Indice : il suffit de montrer que  $\mathbf{F}(x, y)$  n'est pas injective, i.e. qu'il existe deux points différents de  $U$  qui ont la même image.]

d. Soit  $V$  un voisinage de  $\begin{pmatrix} x \\ y \end{pmatrix} \in U$  et soit  $\mathbf{G} : \mathbf{F}(V) \rightarrow V$  un inverse local de  $V$ . Donner la dérivée de  $\mathbf{G}$  en  $\mathbf{F}(x, y)$ .

Essentiellement, le théorème d'inversion locale permet de résoudre l'équation  $\mathbf{x} - \mathbf{F}(\mathbf{y}) = \mathbf{0}$ . Le théorème des applications implicites pousse un peu plus loin en voulant résoudre  $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ , pour  $\mathbf{y}$  en fonction de  $\mathbf{x}$ . Le nom vient du fait que  $\mathbf{y}$  dépend (via  $\mathbf{F}$ ) implicitement de  $\mathbf{x}$ . Dans le cas le plus simple où  $\mathbf{F}$  serait une application linéaire il faut autant d'équations que d'inconnues. Par conséquent, si  $\mathbf{x} \in \mathbb{R}^m$  et  $\mathbf{y} \in \mathbb{R}^n$ ,  $\mathbf{F}$  sera supposée une application de  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  où  $U \subset \mathbb{R}^{m+n}$  est un ouvert.

La notation suivante sera utile pour la démonstration du théorème. Soit  $\mathbf{F} : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$  une application, alors  $\mathbf{F}_{*,[2]}$  sera la matrice  $n \times n$  formée par les  $n$  dernières colonnes de  $\mathbf{F}_*$  (ou autrement dit la matrice de Jacobi de  $\mathbf{F}$  mais en prenant les coefficients  $\mathbf{x}$  comme des constantes), tandis que  $\mathbf{F}_{*,[1]}$  sera la matrice  $n \times m$  formée par les  $m$  premières colonnes de  $\mathbf{F}_*$  (ou autrement dit la matrice de Jacobi de  $\mathbf{F}$  mais en prenant les coefficients  $\mathbf{y}$  comme des constantes).

**Théorème 6.2.5 (Théorème des applications implicites)**

Soit  $U \subset \mathbb{R}^{m+n}$  un ouvert et  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application de classe  $\mathcal{C}^1$  au voisinage de  $\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$  tels que  $\mathbf{F}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$  et  $\text{Det} \mathbf{F}_{*,[2]}(\mathbf{a}, \mathbf{b}) \neq 0$ . Alors  $\exists V_1 \subset \mathbb{R}^m$  un voisinage de  $\mathbf{a}$  et  $V_2 \subset \mathbb{R}^n$  un voisinage de  $\mathbf{b}$  et une application  $\mathbf{T} : V_1 \rightarrow V_2$  de classe  $\mathcal{C}^1$  tels que

$$\forall (\mathbf{x}, \mathbf{y}) \in V_1 \times V_2, \mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0} \Leftrightarrow \mathbf{y} = \mathbf{T}(\mathbf{x}).$$

*Parenthèse :* Le petit frère de ce résultat chez les applications linéaires, est qu'étant données  $n$  équations (linéairement indépendantes) et  $n + m$  variables, alors il y a  $m$  variables "libres" et qui déterminent les  $n$  variables restantes.

DÉMONSTRATION: Soit  $\mathbf{G} : U \rightarrow \mathbb{R}^{m+n}$  l'application définie par  $\mathbf{G}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \mathbf{x} \\ \mathbf{F}(\mathbf{x}, \mathbf{y}) \end{pmatrix}$ . Alors

$$\mathbf{G}_*(\mathbf{x}, \mathbf{y}) = \left( \begin{array}{c|c} \text{Id}_m & \mathbf{0} \\ \hline \mathbf{F}_{*,[1]}(\mathbf{x}, \mathbf{y}) & \mathbf{F}_{*,[2]}(\mathbf{x}, \mathbf{y}) \end{array} \right)$$

Ainsi  $\text{Det} \mathbf{G}_*(\mathbf{a}, \mathbf{b}) = \text{Det} \mathbf{F}_{*,[2]}(\mathbf{a}, \mathbf{b}) \neq 0$ . En particulier le théorème d'inversion locale 6.2.2 implique qu'il existe un voisinage de  $\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$  (dont la forme peut-être supposée  $V_1 \times V_2$ ) tel que  $\mathbf{G}$  possède un inverse  $\mathbf{H}$  de classe  $\mathcal{C}^1$ ,  $\mathbf{H} : \mathbf{G}(V_1 \times V_2) = V_1 \times \mathbf{F}(V_1 \times V_2) \rightarrow V_1 \times V_2$ .

D'un côté ( $\Rightarrow$ ), si  $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in V_1 \times V_2$  est tel que  $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ , alors

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{H} \circ \mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{H}(\mathbf{x}, \mathbf{F}(\mathbf{x}, \mathbf{y})) = \mathbf{H}(\mathbf{x}, \mathbf{0}).$$

Ainsi, en écrivant  $\mathbf{H} = \begin{pmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{pmatrix}$ , il s'obtient que  $\mathbf{y} = \mathbf{H}_2(\mathbf{x}, \mathbf{0})$ . D'où il provient l'application  $\mathbf{T}(\mathbf{x}) = \mathbf{H}_2(\mathbf{x}, \mathbf{0})$  désirée.

De l'autre côté ( $\Leftarrow$ ), si  $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in V_1 \times V_2$  sont tels que  $\mathbf{y} = \mathbf{T}(\mathbf{x}) = \mathbf{H}_2(\mathbf{x}, \mathbf{0})$ , alors

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} = \mathbf{G} \circ \mathbf{H}(\mathbf{x}, \mathbf{0}) = \mathbf{G}(\mathbf{H}_1(\mathbf{x}, \mathbf{0}), \mathbf{H}_2(\mathbf{x}, \mathbf{0})) = \begin{pmatrix} \mathbf{H}_1(\mathbf{x}, \mathbf{0}) \\ \mathbf{F}(\mathbf{H}_1(\mathbf{x}, \mathbf{0}), \mathbf{H}_2(\mathbf{x}, \mathbf{0})) \end{pmatrix} = \begin{pmatrix} \mathbf{H}_1(\mathbf{x}, \mathbf{0}) \\ \mathbf{F}(\mathbf{H}_1(\mathbf{x}, \mathbf{0}), \mathbf{y}) \end{pmatrix}$$

## 6.2 - Théorème des applications implicites

En regardant cette égalité coordonnées par coordonnées, il apparaît que  $\mathbf{x} = \mathbf{H}_1(\mathbf{x}, \mathbf{0})$  et  $\mathbf{0} = \mathbf{F}(\mathbf{H}_1(\mathbf{x}, \mathbf{0}), \mathbf{y})$ . En remplaçant la première équation dans la seconde,  $\mathbf{0} = \mathbf{F}(\mathbf{x}, \mathbf{y})$  ■

Il est facile de montrer que le théorème des fonctions implicites implique le théorème d'inversion locale, il s'agit en réalité de deux énoncés équivalents (le théorème d'inversion locale est le cas particulier où  $\mathbf{F}(\mathbf{x}, \mathbf{y})$  est de la forme  $\mathbf{F}(\mathbf{x}) - \mathbf{y}$ ).

**Exemple 6.2.6.** Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par  $F(x, y) = x^2 + y^2 - 2$  et  $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  alors  $F(1, 1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  et  $F_{*,[2]}(1, 1) = 2 \neq 0$ . Ainsi  $y$  s'exprime localement en fonction de  $x$  au voisinage de  $x = 1$  ; en fait, la résolution directe donne  $y = \sqrt{2 - x^2}$ . Remarquons que si  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  avait été choisi, il aurait fallu prendre  $y = -\sqrt{2 - x^2}$ .

De plus, en  $\begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}$ ,  $F_{*,[2]}(\sqrt{2}, 0) = 0$  et, de fait, il n'est pas possible d'exprimer  $y$  en fonction de  $x$  en ce point. ♣

### 6.2.i Calcul de la dérivée d'une fonction implicite

Même si l'application  $\mathbf{T}$  décrite dans le théorème des applications implicites 6.2.5 paraît très abstraite, il est néanmoins possible de calculer ses dérivées en chaque point. Voici tout d'abord un exemple simple.

**Exemple 6.2.7.** Soit  $F_1, F_2 : \mathbb{R}^4 \rightarrow \mathbb{R}$  deux applications. Supposons qu'il y a des contraintes  $F_1(x_1, x_2, x_3, x_4) = 0$  et  $F_2(x_1, x_2, x_3, x_4) = 0$ . Comment exprimer la variation de  $x_1$  en fonction  $x_3$  en un point  $\mathbf{a} = (a_1, a_2, a_3, a_4)^T$  en considérant  $x_2$  comme étant fixé (dans ce cas  $x_4$  devra aussi varier pour satisfaire les contraintes) ?

Soit  $\mathbf{F}(x_2, x_3, x_1, x_4) = \begin{pmatrix} F_1(x_1, x_2, x_3, x_4) \\ F_2(x_1, x_2, x_3, x_4) \end{pmatrix}$  et supposons alors que  $\text{Det} \mathbf{F}_{*,[2]}(\mathbf{a}) = \frac{\partial F_1}{\partial x_1} \frac{\partial F_2}{\partial x_4} - \frac{\partial F_1}{\partial x_4} \frac{\partial F_2}{\partial x_1} \neq 0$ . Le théorème des fonctions implicites 6.2.5 dit que  $x_1$  et  $x_4$  s'exprime comme des fonctions de  $x_2$  et  $x_3$  au voisinage de  $\mathbf{a}$ .

Pour trouver les dérivées, il n'est pas nécessaire d'inverser la fonction. Il suffit de dériver les deux équations  $F_1 = 0$  et  $F_2 = 0$  par rapport à  $x_3$  :

$$\begin{aligned} \frac{\partial F_1}{\partial x_1} \frac{\partial x_1}{\partial x_3} + \frac{\partial F_1}{\partial x_3} + \frac{\partial F_1}{\partial x_4} \frac{\partial x_4}{\partial x_3} &= 0 \\ \frac{\partial F_2}{\partial x_1} \frac{\partial x_1}{\partial x_3} + \frac{\partial F_2}{\partial x_3} + \frac{\partial F_2}{\partial x_4} \frac{\partial x_4}{\partial x_3} &= 0 \end{aligned}$$

Résoudre (e.g. avec la méthode de Cramer) donne

$$\frac{\partial x_1}{\partial x_3}(a_2, a_3) = \frac{-D_3 F_1(\mathbf{a}) D_4 F_2(\mathbf{a}) + D_4 F_1(\mathbf{a}) D_3 F_2(\mathbf{a})}{D_1 F_1(\mathbf{a}) D_4 F_2(\mathbf{a}) - D_4 F_1(\mathbf{a}) D_1 F_2(\mathbf{a})}$$
 ♣

En général, l'argument se passe de la même façon ( $T_j$  est la  $j^{\text{ème}}$  fonction coordonnée de  $\mathbf{T}$ ) :

#### Proposition 6.2.8

Soit  $U \subset \mathbb{R}^{m+n}$  un ouvert et  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application de classe  $C^1$  au voisinage de  $\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$  tels que  $\mathbf{F}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$  et  $\text{Det} \mathbf{F}_{*,[2]}(\mathbf{a}, \mathbf{b}) \neq 0$ . Soit  $\mathbf{T} : V_1 \rightarrow V_2$  l'application de classe  $C^1$  qui décrit la

dépendance implicite où  $V_1 \subset \mathbb{R}^m$  est un voisinage de  $\mathbf{a}$  et  $V_2 \subset \mathbb{R}^n$  un voisinage de  $\mathbf{b}$ . Alors

$$\forall i \in \overline{m}, \quad \frac{\partial \mathbf{T}}{\partial x_i}(\mathbf{x}) = -\mathbf{F}_{*,[2]}(\mathbf{x}, \mathbf{T}(\mathbf{x}))^{-1} \frac{\partial \mathbf{F}}{\partial x_i}(\mathbf{x}, \mathbf{T}(\mathbf{x})).$$

Dans une écriture moins compacte :

$$\frac{\partial T_j}{\partial x_i}(\mathbf{x}) = -\text{Det} \begin{pmatrix} \frac{\partial \mathbf{F}_1}{\partial y_1} & \frac{\partial \mathbf{F}_1}{\partial y_2} & \dots & \frac{\partial \mathbf{F}_1}{\partial y_{j-1}} & \frac{\partial \mathbf{F}_1}{\partial x_i} & \frac{\partial \mathbf{F}_1}{\partial y_{j+1}} & \dots & \frac{\partial \mathbf{F}_1}{\partial y_n} \\ \frac{\partial \mathbf{F}_2}{\partial y_1} & \frac{\partial \mathbf{F}_2}{\partial y_2} & \dots & \frac{\partial \mathbf{F}_2}{\partial y_{j-1}} & \frac{\partial \mathbf{F}_2}{\partial x_i} & \frac{\partial \mathbf{F}_2}{\partial y_{j+1}} & \dots & \frac{\partial \mathbf{F}_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{F}_n}{\partial y_1} & \frac{\partial \mathbf{F}_n}{\partial y_2} & \dots & \frac{\partial \mathbf{F}_n}{\partial y_{j-1}} & \frac{\partial \mathbf{F}_n}{\partial x_i} & \frac{\partial \mathbf{F}_n}{\partial y_{j+1}} & \dots & \frac{\partial \mathbf{F}_n}{\partial y_n} \end{pmatrix} / \text{Det} \begin{pmatrix} \frac{\partial \mathbf{F}_1}{\partial y_1} & \frac{\partial \mathbf{F}_1}{\partial y_2} & \dots & \frac{\partial \mathbf{F}_1}{\partial y_n} \\ \frac{\partial \mathbf{F}_2}{\partial y_1} & \frac{\partial \mathbf{F}_2}{\partial y_2} & \dots & \frac{\partial \mathbf{F}_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{F}_n}{\partial y_1} & \frac{\partial \mathbf{F}_n}{\partial y_2} & \dots & \frac{\partial \mathbf{F}_n}{\partial y_n} \end{pmatrix}$$

DÉMONSTRATION: L'application  $\mathbf{T}$  est telle que

$$\forall \mathbf{x} \in V_1, \mathbf{F}(\mathbf{x}, \mathbf{T}(\mathbf{x})) = \mathbf{0}.$$

Ainsi, en dérivant par rapport à  $x_i$  (pour  $i \in \overline{m}$ ) et en utilisant la règle de dérivation en chaîne, il apparaît

$$\frac{\partial \mathbf{F}}{\partial x_i}(\mathbf{x}, \mathbf{T}(\mathbf{x})) + \mathbf{F}_{*,[2]}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial \mathbf{T}}{\partial x_i}(\mathbf{x}) = \frac{\partial \mathbf{F}}{\partial x_i}(\mathbf{x}, \mathbf{T}(\mathbf{x})) + \sum_{j \in \overline{n}} \frac{\partial \mathbf{F}}{\partial y_j}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_j}{\partial x_i}(\mathbf{x}) = 0.$$

La formule matricielle apparaît directement. Pour la formulation explicite (et identique), il suffit de résoudre les  $n$  équations (une pour chaque coordonnée de l'application  $\mathbf{F}$ ) du système linéaire :

$$\begin{aligned} \frac{\partial F_1}{\partial y_1}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_1}{\partial x_i}(\mathbf{x}) + \frac{\partial F_1}{\partial y_2}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_2}{\partial x_i}(\mathbf{x}) + \dots + \frac{\partial F_1}{\partial y_n}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_n}{\partial x_i}(\mathbf{x}) &= -\frac{\partial F_1}{\partial x_i}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \\ \frac{\partial F_2}{\partial y_1}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_1}{\partial x_i}(\mathbf{x}) + \frac{\partial F_2}{\partial y_2}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_2}{\partial x_i}(\mathbf{x}) + \dots + \frac{\partial F_2}{\partial y_n}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_n}{\partial x_i}(\mathbf{x}) &= -\frac{\partial F_2}{\partial x_i}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \\ \vdots &+ \vdots + \ddots + \vdots = \vdots \\ \frac{\partial F_n}{\partial y_1}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_1}{\partial x_i}(\mathbf{x}) + \frac{\partial F_n}{\partial y_2}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_2}{\partial x_i}(\mathbf{x}) + \dots + \frac{\partial F_n}{\partial y_n}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \frac{\partial T_n}{\partial x_i}(\mathbf{x}) &= -\frac{\partial F_n}{\partial x_i}(\mathbf{x}, \mathbf{T}(\mathbf{x})) \end{aligned}$$

Ce dernier se résout avec la méthode de Cramer (les inconnus sont ici les  $\frac{\partial T_j}{\partial x_i}$  (où  $j \in \overline{n}$ ), les coefficients les  $\frac{\partial F_k}{\partial y_j}$  (où  $j, k \in \overline{n}$ ) et le vecteur des constantes les  $\frac{\partial F_k}{\partial x_i}$  (où  $k \in \overline{n}$ )) pour conclure la démonstration. ■

EXERCICE 99: Soit  $(x, y, u, v)^T \in \mathbb{R}^4$  tels que

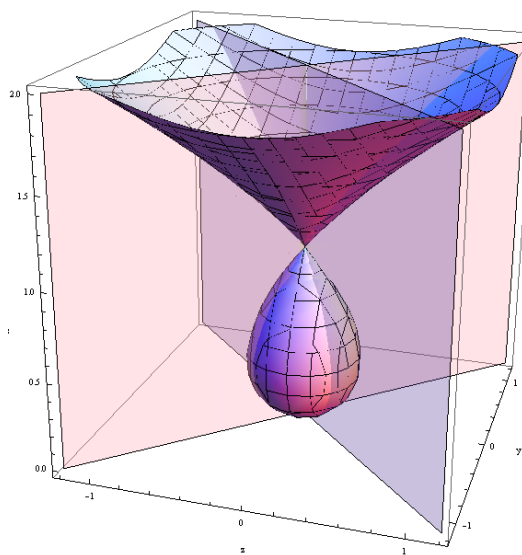
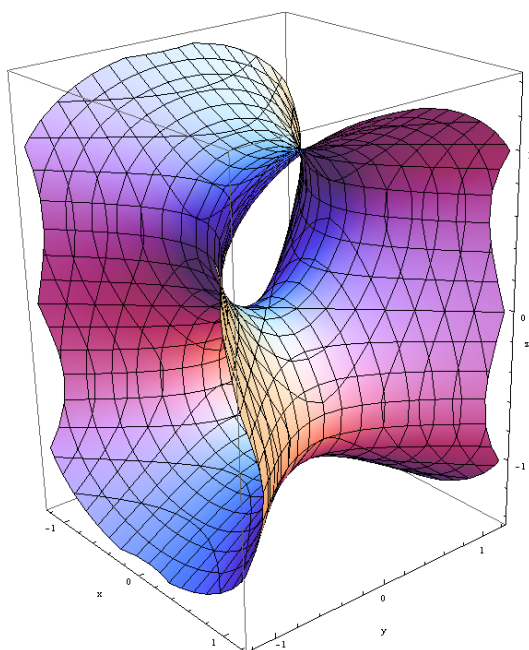
- $x^2 + y^2 + u^2 + v^2 = 1$  et  $x + y + u + v = 0$ ;
- $xyuv = 1$  et  $x + y + u + v = 0$ ,

dans chacun de ces cas dire en quels points  $(p_1, p_2, p_3, p_4)^T$  il est possible d'exprimer (localement !)  $y$  en fonction de  $x$  et  $v$  (c'ad.  $y = y(x, v)$ ), et si oui, déterminer  $\frac{\partial y}{\partial x}$  en ces points (i.e. la variation infinitésimale de  $y$  par rapport à  $x$  en considérant  $v$  fixé).

## 6.2 - Théorème des applications implicites

EXERCICE 100: Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y, z) = x^2 - y^2 + z - \frac{z^3}{3}$ .

- Pourquoi  $F$  est-elle de classe  $C^1$  ?
- Pourquoi l'ensemble des points  $(\mathbf{a}, \mathbf{b}) \in U$  où le théorème des fonctions implicites peut être appliqué est un ouvert ? [Votre réponse devrait rester valable pour n'importe quelle application  $F$  définie d'un ouvert  $U \subset \mathbb{R}^{m+n}$  vers  $\mathbb{R}^n$  qui est de classe  $C^1$ ].
- Au voisinage de quels points  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in F^{-1}(0)$  est-il possible d'exprimer  $z$  comme fonction de  $x$  et  $y$  ? [L'ensemble de niveau  $F^{-1}(0)$  est représenté ci-dessous à gauche.] Soit  $W$  l'ouvert donné par l'ensemble de ces points.
- Soit  $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \in W$  et  $G$  la fonction définie sur  $V \subset \mathbb{R}^2$  un voisinage de  $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$  telle que  $F(x, y, G(x, y)) = 0$  pour tout  $\begin{pmatrix} x \\ y \end{pmatrix} \in V$ . Calculer les dérivées partielles de  $G$  en  $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ .
- Pourquoi est-il ou n'est-il pas possible de définir  $G$  sur un ouvert assez grand de sorte que son image soit tout  $S$  ?
- Pourquoi est-il ou n'est-il pas possible de définir la fonction  $G$  sur  $\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \exists z \in \mathbb{R} \text{ tel que } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in W \}$  ?



EXERCICE 101: Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  l'application définie par  $F(x, y, z) = \begin{pmatrix} z^2 + y^2 - x(x-1)^2 \\ z^2 - y^2 \end{pmatrix}$ . Les ensembles de niveau  $F_1^{-1}(0)$  et  $F_2^{-1}(0)$  sont représentés ci-dessus à droite ; le second est en effet de transparence.

- Au voisinage de quels points  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in F^{-1}(0, 0)$  est-il possible d'exprimer  $y$  et  $z$  comme des fonctions de  $x$  ? Soit  $W$  l'ensemble de ces points.
- $W$  est-il ouvert, fermé, fermé et ouvert, ni fermé ni ouvert ?

- c. Soit  $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \in W$  et  $\gamma$  l'application (une courbe paramétrée) définie sur  $V \subset \mathbb{R}$  un voisinage de  $x_0$  telle que  $F(x, \gamma_1(x), \gamma_2(x)) = \mathbf{0}$  pour tout  $x \in V$ . [Ici,  $\gamma_1$  et  $\gamma_2$  sont les fonctions coordonnées de l'application  $\gamma$ ; ces fonctions sont à valeurs dans  $\mathbb{R}$ .] Calculer la dérivée de  $\gamma$  en  $x_0$ .
- d. Pourquoi est-il ou n'est-il pas possible de définir  $G$  sur un ouvert assez grand de sorte que son image soit tout  $S$  ?

## 6.2.ii Approximation de Taylor pour les fonctions implicites

Grâce au théorème 5.1.7, il est parfois possible de trouver la série de Taylor d'une fonction définie implicitement (sans nécessairement connaître son expression exacte). Seul un exemple sera fait :

**Exemple 6.2.9.** Dans cet exemple les ... seront utilisés pour dire qu'il est toujours possible d'aller chercher des coefficients d'ordre supérieur pour avoir une meilleure information. Soit  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = \sin(x + y) - xy - 2x$ . En  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ,  $F(0, 0) = 0$  et  $F_{*,[2]}(0, 0) = D_2F(0, 0) = \cos 0 - 0 = 1 \neq 0$ . Ainsi dans un voisinage de l'origine,  $y = f(x)$ . Pour trouver les coefficients  $a_i$  de

$$y = f(x) = a_1x + a_2x^2 + a_3x^3 + \dots$$

il est possible de faire un développement de Taylor de

$$\sin(x + f(x)) = \left( (1 + a_1)x + a_2x^2 + a_3x^3 + \dots \right) - \frac{1}{6} \left( (1 + a_1)x + a_2x^2 + a_3x^3 + \dots \right) + \dots$$

et de

$$xf(x) + 2x = 2x + (2 + a_1)x^2 + a_2x^3 + \dots$$

En égalisant ces deux expressions, et, comme elles doivent être égales pour tout  $x$ ,

$$\begin{aligned} 1 + a_1 &= 2 & \Rightarrow a_1 &= 1, \\ a_2 &= a_1 & \Rightarrow a_2 &= 1, \\ a_3 - \frac{1}{6}(1 + a_1)^3 &= a_2 & \Rightarrow a_3 &= \frac{7}{3}. \end{aligned}$$

De la sorte, le développement de Taylor de  $f$  à l'ordre 3 en  $x = 0$  est  $P_{3,0}f(h) = h + h^2 + \frac{7}{3}h^3$  ♣

## 6.3 Espace tangent et sous-variétés

À ce point, trois définitions d'espace tangent ont été introduites, et même s'il y avait une inclusion automatique entre les trois versions, les cas d'égalité sont restés dans l'ombre. Elles sont ici rappelées pour les besoins de la cause. La première est celle en tant qu'ensemble de niveau :

ET1. Soit  $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$  une application de classe  $C^1$  et soit  $S_c = F^{-1}(c)$  pour un  $c \in \mathbb{R}^n$ . Alors  $T_x^{\text{en}} S_c = \text{Ker } F_*(x) = \{v \in \mathbb{R}^m \mid \forall i \in \bar{n}, v \cdot \nabla F_i(x) = 0\}$

En particulier, comme l'espace vectoriel engendré par les  $\nabla F_i$  ( $i \in \bar{n}$ ) est de dimension 0 à  $n$  (inclusivement) l'espace obtenu est de dimension  $m - n$  à  $m$ . Par exemple, si en un  $x$  tous les  $\nabla F_i(x) = 0$ , alors n'importe quel  $v$  appartient à  $T_x S_c$ .



La seconde provient de la notion du vecteur vitesse d'une courbe :

ET2. Soit  $S \subset \mathbb{R}^m$  un ensemble et  $\mathbf{x} \in S$ , alors  $T_{\mathbf{x}}^{vit} S = \{\mathbf{v} \in \mathbb{R}^m \mid \exists \gamma: [-1, 1] \rightarrow \mathbb{R}^m \text{ telle que } \gamma \text{ est de classe } C^1, \gamma([-1, 1]) \subset S, \gamma(0) = \mathbf{x} \text{ et } \gamma'(0) = \mathbf{v}\}$ .

Autrement dit, c'est l'ensemble des vecteurs basés en  $\mathbf{x}$  qui peuvent être réalisés comme le vecteur vitesse d'une courbe. **Attention** : il ne s'agit en général pas d'un espace vectoriel (seulement d'un espace homogène).

La dernière provient de la notion de paramétrage :

ET3. Soit  $S \subset \mathbb{R}^m$  un ensemble,  $\mathbf{G}: \mathbb{R}^{\ell} \rightarrow S$  un paramétrage de classe  $C^1$  de  $S$ , alors  $T_{\mathbf{y}}^{par} \mathbf{G} = \text{Im } \mathbf{G}_*(\mathbf{y}) = \{\mathbf{v} \in \mathbb{R}^m \mid \mathbf{v} \text{ est une combinaison linéaire des } D_j \mathbf{G}(\mathbf{y}), \text{ où } j \in \bar{\ell}\}$ .

L'espace tangent ci-dessus a quant à lui une dimension incluse entre 0 et  $\ell$  (puisque'il est engendré par  $\ell$  vecteurs). Les cas où  $T^{par}$  (resp.  $T^{en}$ ) est de dimension maximale (resp. minimale) sont très importants, voir le corollaire 3.7.10. Ils se verront à partir de maintenant attribuer un nom.

Avant de poursuivre, il est bon de relire la section 3.7 ou au moins la sous-section 3.7.iv. En bref, les trois définitions de plan tangent forment une inclusion évidente :  $T^{par} \subset T^{vit} \subset T^{en}$ . Celui du milieu n'est, en général, pas un espace vectoriel. Mais lorsque les deux du bout sont de même dimension, il en devient automatiquement égal à ces deux espaces et il devient lui-même automatiquement un espace vectoriel (cf. corollaire 3.7.10). Des cas où  $T^{vit}$  n'est pas un espace vectoriel ont déjà été abordés (cf. exercice 49) ; il arrive même que les trois espaces tangents soient des espaces vectoriels distincts :

**Exemple 6.3.1.** Soit  $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 = z^3\}$ . Bien que légèrement compliqué, il est possible de montrer que  $T_0^{vit} S$  est l'espace vectoriel engendré par le vecteur  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  (c'est cependant facile de voir qu'il contient cet espace car la courbe paramétrée  $\gamma(t) = \begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix}$  est d'image contenue dans  $S$ ).

Maintenant,  $S = S_0$  pour la fonction  $F: \mathbb{R}^3 \rightarrow \mathbb{R}$  définie par  $F(x, y, z) = z^3 - x^2$ . Or  $\nabla F(0, 0, 0) = \mathbf{0}$ , et ainsi  $T_0^{en} S_0 = \mathbb{R}^3$ .

Finalement, l'application  $\mathbf{G}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  définie par  $\mathbf{G}(s, t) = \begin{pmatrix} s^3 \\ t^3 \\ s^2 \end{pmatrix}$  est  $C^1$  et même une bijection entre  $\mathbb{R}^2$  et  $S$ . Cependant,  $T_0^{par} \mathbf{G} = \{\mathbf{0}\}$  (ici  $\mathbf{G}(\mathbf{0}) = \mathbf{0}$  donc c'est bien ce point du paramétrage qu'il faut regarder). ♣

Quelle que soit la méthode pour définir un sous-ensemble de  $\mathbb{R}^m$  (i.e. comme ensemble de niveau dans ET1, comme image d'une application dans ET3, ou tout simplement comme sous-ensemble dans ET2), un point est singulier lorsque les trois définitions de plan tangent ne peuvent coïncider.

En réalité, le seul plan tangent qui nous intéresse vraiment, est celui des vecteurs vitesses : c'est la seule définition "intrinsèque" (i.e. propre à l'ensemble). Comme on a vu, il est difficile à calculer tandis que ses deux "cousins"  $T^{par}$  et  $T^{en}$  sont faciles à calculer. Une des applications majeure du théorème d'inversion locale et du théorème des fonctions implicites est de donner des critères qui permettent d'assurer qu'une des notions "calculables" ( $T^{par}$  ou  $T^{en}$ ) est égale au plan tangent  $T^{vit}$ .

Le résultat sera même plus fort que celui demandé, au sens où on montrera que les points singuliers forment un fermé. En particulier, une petite perturbation des conditions fera qu'on se trouve dans une situation régulière. Voici les définitions clefs de cette section :



**Définition 6.3.2.** Soit  $U \subset \mathbb{R}^\ell$  un ouvert et  $\mathbf{G} : U \rightarrow S \subset \mathbb{R}^m$  une application.  $\mathbf{y} \in U$  est dit un point régulier (du paramétrage) si une des conditions équivalentes suivante est satisfaite :

- $\dim T_{\mathbf{y}}^{\text{par}} \mathbf{G} = \ell$  ;
- les  $D_i \mathbf{G}(\mathbf{y})$  (où  $i \in \bar{\ell}$ ) sont linéairement indépendants ;
- la matrice  $\mathbf{G}_*(\mathbf{y})$  possède une sous-matrice de taille  $\ell \times \ell$  qui est de déterminant non-nul.

Si un paramétrage est injectif et que tous les points sont réguliers, il sera dit un **paramétrage local**. ★

Des définitions similaires peuvent être faites pour les ensembles de niveaux :

**Définition 6.3.3.** Soit  $V \subset \mathbb{R}^m$  et  $\mathbf{F} : V \rightarrow \mathbb{R}^n$  une application de type  $C^1$ . Le point  $\mathbf{x} \in S_{\mathbf{c}} := \mathbf{F}^{-1}(\mathbf{c}) \subset \mathbb{R}^m$  est dit un point régulier (de l'ensemble de niveau) si une des conditions équivalentes suivante est satisfaite :

- $\forall \mathbf{x} \in S_{\mathbf{c}}, \dim T_{\mathbf{x}}^{\text{en}} S_{\mathbf{c}} = m - n$  ;
- les gradients  $\nabla F_j(\mathbf{x})$  (où  $j \in \bar{n}$ ) sont linéairement indépendants ;
- la matrice  $\mathbf{F}_*(\mathbf{y})$  possède une sous-matrice de taille  $m \times m$  qui est de déterminant non-nul.

La valeur  $\mathbf{c} \in \mathbb{R}^n$  est dite une valeur régulière de  $\mathbf{F}$  si tout  $\mathbf{x} \in S_{\mathbf{c}} := \mathbf{F}^{-1}(\mathbf{c})$  est un point régulier. ★

Si la terminologie ci-dessus est redondante, c'est que l'objectif de cette section est précisément de montrer que les points réguliers d'un paramétrage sont des points réguliers d'un ensemble de niveau et *vice versa*.

Dans l'énoncé suivant, lorsque  $\ell = m$ ,  $\mathbb{R}^{m-\ell}$  sera entendu comme étant l'espace vectoriel de dimension 0 :  $\{0\}$ .

#### Théorème 6.3.4

Soit  $\ell \leq m \in \mathbb{Z}_{\geq 1}$ , soit  $V \subset \mathbb{R}^\ell$  un ouvert,  $\mathbf{G} : V \rightarrow \mathbb{R}^m$  une application de classe  $C^1$ , et  $\mathbf{b} \in V$  un point régulier de  $\mathbf{G}$ . Alors il existe un voisinage  $U \subset V \subset \mathbb{R}^\ell$  de  $\mathbf{b}$ ,  $W$  un ouvert contenant  $\mathbf{G}(U)$  et une application  $\mathbf{F} : W \rightarrow \mathbb{R}^{m-\ell}$  telle que, en posant  $\mathbf{a} := \mathbf{G}(\mathbf{b})$ ,

- $\mathbf{G}(U) = \mathbf{F}^{-1}(\mathbf{0}) =: S_{\mathbf{0}}$ ,
- $T_{\mathbf{b}}^{\text{par}} \mathbf{G} = T_{\mathbf{a}}^{\text{en}} S_{\mathbf{0}}$
- et  $\mathbf{0}$  est une valeur régulière de  $\mathbf{F}$ .

En particulier, si  $S = \mathbf{G}(U)$ ,  $T_{\mathbf{a}}^{\text{vit}} S$  est égal aux deux autres plan tangents ( $T_{\mathbf{b}}^{\text{par}} \mathbf{G}$  et  $T_{\mathbf{a}}^{\text{en}} S_{\mathbf{0}}$  (par la proposition 3.7.9).

**DÉMONSTRATION:** Grâce à l'indépendance linéaire, une des sous-matrices carrées  $\ell \times \ell$  de  $\mathbf{G}_*(\mathbf{b})$  aura un déterminant non-nul, et (en réordonnant les coordonnées) il est possible de supposer qu'il s'agit des  $\ell$  premières colonnes. Par conséquent, l'application  $\mathbf{H}(\mathbf{y}) = \begin{pmatrix} G_1(\mathbf{y}) \\ \vdots \\ G_\ell(\mathbf{y}) \end{pmatrix}$  de  $V$  dans  $\mathbb{R}^\ell$  est, sur un voisinage  $U$  de  $\mathbf{b}$ , un difféomorphisme  $C^1$  (par le théorème d'inversion locale 6.2.2). En mots,  $\mathbf{H}^{-1}$  associe aux  $\ell$  premières coordonnées d'un point de  $\text{Im } \mathbf{G}$ , l'unique élément de  $V$  qui est envoyé dessus.

Si  $m = \ell$ ,  $\mathbf{G}$  est donc surjective sur n'importe quel voisinage (dans  $\mathbb{R}^m$ ) d'un point de  $\mathbf{G}(U)$  ainsi la fonction  $F$  qui envoie tous les points vers  $\{0\}$  fait l'affaire.

Si  $m > \ell$ , soit  $\mathbf{x} \in \mathbf{G}(U)$ , alors

$$\mathbf{G} \circ \mathbf{H}^{-1}(x_1, x_2, \dots, x_\ell) = \begin{pmatrix} x_1 \\ \vdots \\ x_\ell \\ G_{\ell+1}(\mathbf{H}^{-1}(x_1, x_2, \dots, x_\ell)) \\ \vdots \\ G_m(\mathbf{H}^{-1}(x_1, x_2, \dots, x_\ell)) \end{pmatrix}$$

Par conséquent,  $\mathbf{F}$  définie sur  $W = \text{Im } \mathbf{H} \times \mathbb{R}^{m-\ell}$  par  $F_i(\mathbf{x}) = G_{i+\ell}(x_1, x_2, \dots, x_\ell) - x_{i+\ell}$ , où  $i \in \overline{m-\ell}$ , est telle que  $\mathbf{F}^{-1}(\mathbf{0}) = \text{Im } \mathbf{G}$ .

Il ne reste plus qu'à remarquer que les  $\nabla F_i(\mathbf{a})$  sont linéairement indépendants (les  $\ell$  premières coordonnées de chaque gradient sont un peu étranges, bien que calculables, et les  $m - \ell$  dernières coordonnées sont des 0 et un 1). Ainsi  $\dim T_{\mathbf{a}}^{\text{en}} S_0$  est l'orthogonale dans  $\mathbb{R}^m$  d'un espace de dimension  $m - \ell$ , soit un espace vectoriel de dimension  $\ell$ . D'autre part, comme les  $\ell$  colonnes de  $\mathbf{G}_*(\mathbf{y})$  (i.e. les  $D_i \mathbf{G}(\mathbf{y})$ ) sont linéairement indépendantes, la dimension de  $T_{\mathbf{b}}^{\text{par}} \mathbf{G}$  est aussi  $\ell$ . Ainsi ces deux espaces vectoriels doivent être identiques. ■

Le précédent théorème dit en fait plus que ce qui avait été annoncé : si le paramétrage est bon, il est toujours (localement !) possible de réaliser l'image comme un ensemble de niveau et les plans tangents sont indépendants du point de vue. Le prochain donnera le sens inverse, c'ad. que si un ensemble de niveau est "régulier" il peut être (localement !) vu comme l'image d'un paramétrage et les notions de plan tangent coïncident toutes.

### Théorème 6.3.5

Soit  $m \geq n \in \mathbb{Z}_{\geq 1}$ ,  $U$  un ouvert de  $\mathbb{R}^m$  et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  une application de classe  $\mathcal{C}^1$ , soit  $S = S_0 = \mathbf{F}^{-1}(\mathbf{0})$ . Soit  $\mathbf{a} \in U$  tel que  $\dim T_{\mathbf{a}}^{\text{en}} S_0 = m - n$  alors il existe un voisinage  $V$  de  $\mathbf{a}$ ,  $W \subset \mathbb{R}^{m-n}$  un ouvert et  $\mathbf{G} : W \rightarrow S_0 \cap V \subset \mathbb{R}^m$  une application telle que

- $\mathbf{G}$  est bijective et  $\mathcal{C}^1$  ;
- si  $\mathbf{b} = \mathbf{G}^{-1}(\mathbf{a})$ ,  $T_{\mathbf{b}}^{\text{par}} \mathbf{G} = T_{\mathbf{a}}^{\text{en}} S_0$  ;
- et tout point de  $W$  est un point régulier de  $\mathbf{G}$ .

DÉMONSTRATION: Si  $m = n$  alors  $\mathbf{F}$  est en fait un difféomorphisme local en  $\mathbf{x}$ , ainsi  $\mathbf{F}^{-1}(\mathbf{0}) = \{\mathbf{x}\}$  et  $\mathbf{G} : \{0\} \rightarrow \{\mathbf{x}\}$  ne peut être définie que d'une façon.

Lorsque  $m > n$ , c'est le théorème des fonctions implicites 6.2.5 qui sera utilisé (rappelons qu'il est équivalent au théorème d'inversion locale 6.2.2). En effet, comme les gradients sont linéairement indépendants (et supposons que  $m > n$ ) alors une sous-matrice  $n \times n$  de  $\mathbf{F}_*(\mathbf{a})$  sera de déterminant non-nul. Quitte à réordonner les coordonnées, il sera supposé qu'il s'agit de celle formée par les  $n$  dernières colonnes. Posons  $\ell = m - n$ ,  $\mathbf{F}$  va d'un ouvert de  $\mathbb{R}^{\ell+n}$  dans un ouvert de  $\mathbb{R}^n$ . Il est alors possible d'appliquer le théorème des fonctions implicites 6.2.5 pour montrer qu'il existe un voisinage  $V_1 \times V_2$  de  $\mathbf{a}$  (où  $V_1 \subset \mathbb{R}^\ell$  et  $V_2 \subset \mathbb{R}^n$ ) et  $\mathbf{T} : V_1 \rightarrow V_2$  de classe  $\mathcal{C}^1$  tel que

$$\forall \mathbf{x} \in V_1 \times V_2, \quad F(\mathbf{x}) = \mathbf{0} \Leftrightarrow (x_{\ell+1}, \dots, x_m)^\top = \mathbf{T}(x_1, \dots, x_\ell).$$

C'est essentiellement le paramétrage : soit  $\mathbf{G} : V_1 \rightarrow V_1 \times V_2$  défini par  $\mathbf{G}(\mathbf{y}) = (\text{Id}_{\ell \times \ell}(\mathbf{y}), \mathbf{T}(\mathbf{y}))$ . Alors par construction,  $\mathbf{G}$  est une application bijective et  $\mathcal{C}^1$  à valeur dans  $\mathbf{F}^{-1}(\mathbf{0})$ .

Pour montrer l'égalité des espaces tangents il suffit de voir que la dimension est égale. L'espace tangent  $T_{\mathbf{a}}^{\text{en}} S_0$  est donné par orthogonalité à  $n$  vecteurs linéairement indépendants ; sa dimension est  $m - n = \ell$ . Quant à  $T_{\mathbf{b}}^{\text{par}} \mathbf{G}$ , comme  $\mathbf{G}_*(\mathbf{b})$  contient une sous-matrice de la forme identité, les  $D_i \mathbf{G}(\mathbf{b})$  sont linéairement indépendants ; sa dimension est aussi  $\ell = m - n$ . ■

Les théorèmes 6.3.4 et 6.3.5 ont deux conséquences importantes à souligner :

- Le point régulier d'un paramétrage est localement un point régulier d'un ensemble de niveau. *Vice-versa*, le point régulier d'un ensemble de niveau est localement un point régulier d'un paramétrage
- Tout point régulier possède un voisinage formé de points réguliers, *i.e.* l'ensemble des points réguliers est un ouvert (pour la topologie induite sur  $S$ ).

Le concept de variété sera surtout important lorsque les surfaces seront abordées, mais il est important de voir que c'est aussi la condition qui fait que la méthode des multiplicateurs de Lagrange puisse être appliquée.

**Définition 6.3.6.** Soit  $S$  un sous-ensemble de  $\mathbb{R}^m$ .  $S$  est une [sous-]variété de  $\mathbb{R}^m$ , si une des deux conditions équivalentes<sup>2</sup> suivantes est vérifiée :

- $\forall \mathbf{x} \in S$  il existe un voisinage  $U$  de  $\mathbf{x}$  (dans  $\mathbb{R}^m$ ) et un paramétrage local  $\mathbf{G} : V \rightarrow U \cap S$  où  $V \subset \mathbb{R}^\ell$  est un ouvert.
- $\forall \mathbf{x} \in S$  il existe un voisinage  $U$  de  $\mathbf{x}$  (dans  $\mathbb{R}^m$ ) et une application  $\mathbf{F} : U \rightarrow \mathbb{R}^n$  telle que  $S = \mathbf{F}^{-1}(\mathbf{0})$  et  $\mathbf{0}$  est une valeur régulière de  $\mathbf{F}$ .

La **dimension** d'une sous-variété connexe  $S$  de  $\mathbb{R}^m$  est, dans le premier cas, l'entier  $\ell$ , et, dans le second cas, l'entier  $m - n$ . ★

Il est bon de noter que les entiers  $m - n$  et  $\ell$  sont tout à fait en analogie avec les dimensions de noyau et d'image en algèbre linéaire. Un paramétrage local de  $S$  est dit paramétrage global s'il est surjectif sur  $S$  (et donc bijectif).

Il serait tentant de croire qu'une des deux perceptions d'un sous-ensemble (via les ensembles de niveaux ou via les paramétrages locaux) est meilleure que l'autre. Au niveau local, les théorèmes 6.3.4 et 6.3.5 montrent qu'il n'y a aucune différence. Voici une liste d'exemples qui indique que dans le cas global, il n'y a pas de règle :

**Exemple 6.3.7.** La sphère ne peut pas être l'image d'un seul paramétrage local, mais elle est définie par une surface de niveau :  $x^2 + y^2 + z^2 = 1$ . Elle peut être donnée par l'image d'une application (une surface paramétrée) régulière, mais celle-ci n'est pas injective. Deux paramétrages locaux possibles sur la sphère sont les deux projections stéréographiques.

L'hyperboloïde à deux nappes est une surface de niveau correspondant à l'équation  $z^2 = 1 + x^2 + y^2$ . Il est possible de se restreindre à une seule de ses "nappes", en regardant par exemple  $S_0 = F^{-1}(0)$  où  $F(x, y, z) = 1 + x^2 + y^2 - z^2$  est définie non pas sur  $\mathbb{R}^3$  mais sur  $\mathbb{R} \times \mathbb{R} \times \mathbb{R}_{\geq 0}$ . Alors,

2. Cette équivalence est le contenu des théorèmes 6.3.5 et 6.3.4.

cet ensemble de niveau possède un paramétrage global (il suffit de projeter sur le plan  $xy$  et de prendre l'inverse de cette projection :  $\mathbf{G}(x,y) = \begin{pmatrix} x \\ y \\ \sqrt{1+x^2+y^2} \end{pmatrix}$ ).

L'hyperboloïde à une nappe est aussi une surface de niveau :  $x^2 + y^2 - z^2 = 1$ . Il est possible de le décrire en utilisant un paramétrage global (il faut chercher un peu). Il est par contre l'image d'une application (surface paramétrée) régulière (cf. exemple 8.0.2) non-injective (ainsi deux paramétrages locaux qui recouvrent cette nappe peuvent être formés par deux restrictions de ce paramétrage sur deux domaines différents).

Le ruban de Möbius (voir l'exemple 8.2.5) ne peut ni être couvert par un paramétrage global ni être décrit comme une surface de niveau (combinaison du fait qu'il n'est pas orientable avec la proposition 8.1.2). Par contre, il est néanmoins possible de le décrire comme l'image d'une application (surface paramétrée) régulière (non-injective, bien sûr). ♣

EXERCICE 102: Montrer que le (double-)cône  $S = \{(x,y,z)^T \in \mathbb{R}^3 \mid x^2 + y^2 = z^2\}$  n'est pas une sous-variété de  $\mathbb{R}^3$ .

[Attention : il est insuffisant de montrer que le gradient d'une fonction dont  $S$  est une surface de niveau est nul, il faudrait le montrer pour toutes les fonctions qui réalisent  $S$  comme surface de niveau. Par contre, il est plus facile (et très recommandé) de montrer que  $T_0^{vit} S$  n'est pas un espace vectoriel.]

Avant de clore cette section sur l'espace tangent, voici un exercice qui devrait permettre de sceller l'interprétation de cet espace vectoriel. Un court rappel est de bon ton. Soit  $S \subset \mathbb{R}^m$  une sous-variété. Pour rappel, étant donné une des trois définitions d'espace tangent  $T^{par}$ ,  $T^{vit}$  ou  $T^{en}$  (qui coïncident toutes comme  $S$  est une sous-variété), il est toujours possible de les réaliser non pas comme un espace vectoriel mais comme l'ensemble (affin, e.g. en dimension 2 le plan) qui intersecte  $S$  en (au moins) un point. Par exemple, le "plan" tangent (en tant qu'en ensemble) en  $\mathbf{a} \in S$  est défini par  $P_{\mathbf{a}}^{vit} = \{\mathbf{a} + \mathbf{v} \mid \mathbf{v} \in T_{\mathbf{a}}^{vit} S\}$ .

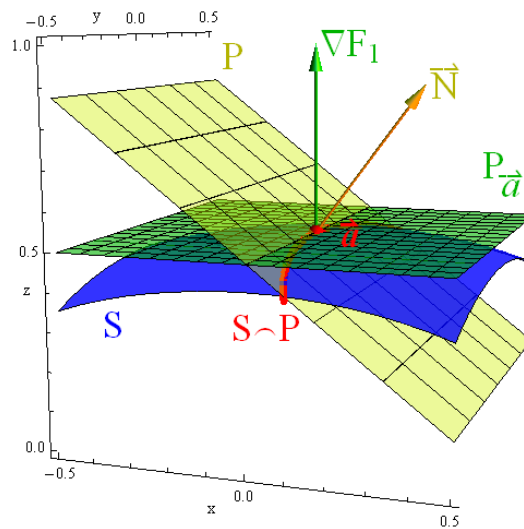
EXERCICE 103: Soit  $S \subset \mathbb{R}^m$  une sous-variété de dimension  $\ell \in \mathbb{Z}_{\geq 0}$  tel que  $1 < \ell < m$ . Montrer que si  $P$  est un hyperplan de  $\mathbb{R}^m$  (i.e. un ensemble où une fonction affine s'annule) tel que  $P \cap S$  est exactement un point  $\mathbf{a}$ , alors  $P$  contient le plan tangent de  $P_{\mathbf{a}} S$ .

[Indices : Ci-dessous le dessin d'un cas où le vecteur normal est linéairement indépendant du gradient.]

1- Remarquer que si  $A = \{\mathbf{a}\}$  est un point, alors  $T_{\mathbf{a}}^{vit} A = \{\mathbf{0}\}$ .

2- Réaliser  $S$  localement autour de  $\mathbf{a}$  comme l'ensemble de niveau d'une application  $\mathbf{F}$  et réaliser  $P$  (globalement) comme l'ensemble de niveau d'une fonction  $A$  (affine). En déduire que  $S \cap P$  est aussi localement l'ensemble de niveau d'une application.

3- Montrer que si le vecteur  $\mathbf{v}$  normal à  $P$  (i.e. tel que  $P = \{\mathbf{a} + \mathbf{w} \mid \mathbf{w} \cdot \mathbf{v} = 0\}$ ) est linéairement indépendant l'espace engendré par les  $\nabla F_j(\mathbf{a})$  (où  $j \in \overline{m-\ell}$ ) alors le plan tangent en  $\mathbf{a}$  est de dimension (strictement) supérieure à 0. ]



EXERCICE 104: Soit  $\text{On}(\mathbb{R})$  l'ensemble des matrices orthogonales à coefficients réels, i.e. le sous-ensemble de  $\mathcal{M}_{n \times n}(\mathbb{R})$  formé des matrices  $A$  telles que  $A^T A = \text{Id}$ . L'objectif est de montrer que ces matrices forment une sous-variété de  $\mathcal{M}_{n \times n}(\mathbb{R})$ . Pour ce faire, il faut voir  $\mathcal{M}_{n \times n}(\mathbb{R})$  comme  $\mathbb{R}^{n^2}$  (puisque'il y a  $n^2$  coefficients dans ces matrices et qu'elles forment un espace vectoriel (via l'addition et la multiplication par un scalaire) et même euclidien (via la norme  $\|\cdot\|_2$ )).

- Soit  $S$  le sous-espace vectoriel formé des matrices symétriques de taille  $n \times n$ . Dire pourquoi  $S$  est un espace vectoriel de dimension  $n(n+1)/2$ .
- Soit  $F : \mathcal{M}_{n \times n} \rightarrow S$  la fonction définie par  $F(A) = A^T A$ . Pour  $B \in \mathcal{M}_{n \times n}$ , calculer la dérivée directionnelle

$$D_B F(A) = \lim_{t \rightarrow 0} \frac{F(A + tB) - F(A)}{t}$$

- Montrer que, si  $B$  est fixée  $D_B F(A)$  est une application linéaire, i.e.  $D_B F(A) + D_B F(A') = D_B F(A + A')$ . En déduire que  $F$  est de classe  $C^1$  (inutile de prendre directement la définition de limite ; seulement les propriétés sur la somme/composition/produit de fonctions  $C^1$ ).
- Il faut montrer que  $\text{Id}$  est une valeur régulière de  $F$ . Pourquoi est-il suffisant de montrer que

$$\forall A \in F^{-1}(\text{Id}), \forall M \in S, \exists B \in \mathcal{M}_{n \times n} \text{ tel que } D_B F(A) = M.$$

- Montrer que  $\text{Id}$  est une valeur régulière de  $F$ , i.e. que  $\text{O}(n)$  est une sous-variété. Quelle est la dimension de  $\text{O}(n)$  ?
- Montrer que son espace tangent en  $A = \text{Id}$  est formé par les matrices antisymétriques, i.e. montrer que  $D_B F(\text{Id}) = 0$  implique que  $B^T = -B$ .

EXERCICE 105: Montrer que le graphe d'une fonction est une sous-variété. Pour ce faire, il est possible de regarder l'exercice 54 (mais avec les théorèmes maintenant à disposition, il n'est pas nécessaire de refaire tout l'exercice).

### 6.3.i Cas particuliers

Le chapitre 4 a déjà traité les variétés les plus simples. Cependant, les courbes ne sont pas toutes des variétés : si il y a un point de croisement<sup>3</sup> ou des point de “rebroussement”<sup>4</sup>.

Par contre, une courbe régulière et injective aura toujours pour image une variété de dimension 1. D’ailleurs, le mot régulier est antonyme de singulier et son origine est liée qu fait que les points où le plan tangent des vecteurs vitesse ne peut pas correspondre à un plan tangent “simple” (paramétrage ou ensemble de niveau) était autrefois appelé des points singuliers ou singularités.

Le chapitre 10 traitera quant à lui des surfaces paramétrées qui seront, lorsque des hypothèses de régularité et d’injectivité sont satisfaites, des variétés de dimension 2.

## 6.4 Multiplicateurs de Lagrange

Une question qui survient tout au moins aussi souvent que “chercher un extremum d’une fonction” est celle de trouver un extremum sous certaines contraintes. Par surcroît, cette idée est aussi utile pour trouver les extrema absolus (s’ils existent).

Une contrainte sera ici de ne pas considérer tous les  $\mathbf{x}$  dans le domaine de définition  $U$  de la fonction  $F$  mais seulement les  $\mathbf{x}$  qui appartiennent à un sous ensemble  $S \subset U$ . Afin de pouvoir traiter cette nouvelle contrainte (et comme ces contraintes prennent habituellement la forme d’une équation) il sera supposé que  $S = \mathbf{G}^{-1}(\mathbf{0})$  où  $\mathbf{G} : U \rightarrow \mathbb{R}^n$  est une application telle que  $U$  est un ouvert contenant  $S$ , c’ad. que  $S$  est un ensemble de niveau pour une application  $\mathbf{G}$  (cf. sections 3.7 et 6.3). Afin de pouvoir être traité par le biais du calcul différentiel,  $\mathbf{G}$  sera elle-même supposée être de classe  $\mathcal{C}^1$  sauf peut-être en quelques points.

Une autre façon, équivalente, de présenter la chose, est de dire que seuls les  $\mathbf{x}$  satisfaisant  $\mathbf{G}(\mathbf{x}) = \mathbf{0}$  seront considérés. Si, pour  $j \in \overline{n}$ ,  $G_j$  est la  $j^{\text{ème}}$  fonction coordonnée de  $\mathbf{G}$ , il est utile de rappeler que  $\mathbf{G}^{-1}(\mathbf{0}) = \cap_{j \in \overline{n}} G_j^{-1}(0)$ . La question est formulée par la définition suivante :

**Définition 6.4.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert et soit  $F : U \rightarrow \mathbb{R}$  une fonction. Soit  $S := \mathbf{G}^{-1}(\mathbf{0}) \subset U$ , où  $\mathbf{G} : U \rightarrow \mathbb{R}^n$  est une application, la contrainte. Le point  $\mathbf{x} \in S$  est un maximum (resp. minimum) local de  $F$  sous la contrainte  $S$  si  $\exists r \in \mathbb{R}_{>0} \forall \mathbf{y} \in B_r(\mathbf{x}) \cap S, F(\mathbf{y}) < F(\mathbf{x})$  (resp.  $>$ ). ★

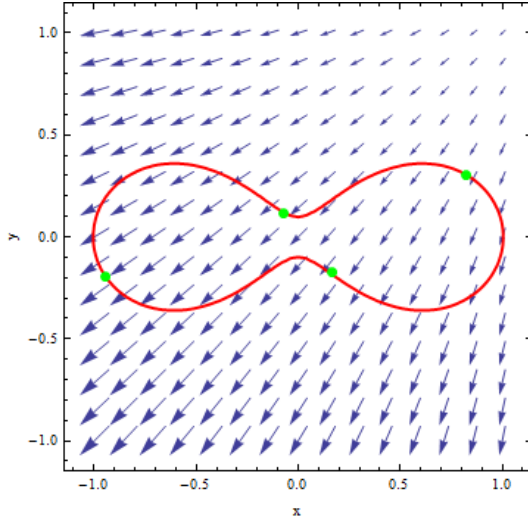
Il ne s’agit en quelque sorte que de chercher les extrema de la fonction  $F$  restreinte à  $S$ ,  $F|_S : U \cap S \rightarrow \mathbb{R}$ .

Afin de présenter l’idée de la réponse,  $\mathbf{G}$  sera temporairement supposée être une fonction ; la contrainte est temporairement réduite  $G_1(\mathbf{x}) = 0$ . D’un côté, le gradient de  $G_1$  est perpendiculaire à l’ensemble de niveau  $G_1(\mathbf{x}) = 0$ . D’un autre, le gradient de  $F$  indique la variation la plus importante de  $F$ . En un point  $\mathbf{x}$  de  $U \cap S$ , si  $\nabla F(\mathbf{x})$  n’est pas perpendiculaire à l’ensemble de niveau, sa composante tangente à l’ensemble de niveau  $\mathbf{v} := \nabla F(\mathbf{x}) (1 - \frac{\nabla F(\mathbf{x}) \cdot \nabla G_1(\mathbf{x})}{\|\nabla F(\mathbf{x})\| \|\nabla G_1(\mathbf{x})\|})$  est non nulle. Pour

3. Puisque la variété aura alors essentiellement l’air de l’ensemble de niveau  $\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 = y^2 \}$  où  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est le point de croisement.

4. Comme dans la parabole semi-cubique, l’ensemble de niveau  $\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 = y^3 \}$  ou l’image de la courbe  $\gamma(t) = \begin{pmatrix} t^3 \\ t^2 \end{pmatrix}$  où, dans les deux cas,  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  est le point de rebroussement.

$t$  petit  $\mathbf{x} + t\mathbf{v}$ , donnera une variation stricte (du même signe que  $t$ ) de la valeur de  $F$  ET reste très proche de la contrainte :  $G_1(\mathbf{x} + t\mathbf{v})$  est proche de 0 (puisque  $\mathbf{v}$  est tangent à la surface de niveau). En déformant un peu, il est raisonnable de croire que des points proprement dans  $S$  auront des valeurs strictement plus grandes et plus petites. La condition est donc que si  $\mathbf{x}$  est un extrema,  $\nabla F$  doit être colinéaire à  $\nabla G_1$ .



Par exemple, dans le dessin ci-contre est représenté le gradient d'une fonction (par les flèches) et un ensemble de niveau. Les points où la fonction prend sa plus petite et plus grande valeur sont parmi les quatre points où le gradient est orthogonal au plan tangent à l'ensemble de niveau.

La restriction à des ensembles de niveaux de la forme  $\mathbf{G}^{-1}(\mathbf{0})$  plutôt que  $\mathbf{G}^{-1}(\mathbf{v})$  pour  $\mathbf{v} \in \mathbb{R}^n$  n'est qu'apparente : il suffit de définir  $\tilde{\mathbf{G}}(\mathbf{x}) = \mathbf{G}(\mathbf{x}) - \mathbf{v}$  pour passer de l'un à l'autre :  $\tilde{\mathbf{G}}^{-1}(\mathbf{0}) = \mathbf{G}^{-1}(\mathbf{v})$ .

#### Théorème 6.4.2

Soit  $U \subset \mathbb{R}^m$  un ouvert, soit  $F : U \rightarrow \mathbb{R}$  une fonction dérivable et soit  $\mathbf{G} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  une application de classe  $C^1$  telle que  $S := \mathbf{G}^{-1}(\mathbf{0}) \neq \emptyset$  et 0 est une valeur régulière<sup>5</sup> de  $\mathbf{G}$ . Si  $\mathbf{x}$  est un extremum local de  $F|_S$  alors  $\exists \mathbf{a} \in \mathbb{R}^n$  tel que

$$\nabla F(\mathbf{x}) = \sum_{i=1}^n a_i \nabla G_i(\mathbf{x}).$$

DÉMONSTRATION: Étant donné que  $S$  est une valeur régulière de  $\mathbf{G}$ ,  $T_{\mathbf{x}}^{\text{vir}} S = T_{\mathbf{x}}^{\text{en}} S$ . Autrement dit, pour tout  $\mathbf{v} \in \mathbb{R}^m$  un vecteur qui est orthogonal à tous les  $\nabla G_i(\mathbf{x})$  (où  $i \in \overline{n}$ ) il existe une courbe paramétrée  $\gamma(t) : [-1, 1] \rightarrow \mathbb{R}^m$  de classe  $C^1$  telle que  $\gamma([-1, 1]) \subset S = \mathbf{G}^{-1}(\mathbf{0})$ ,  $\gamma(0) = \mathbf{x}$  et  $\gamma'(0) = \mathbf{v}$ . Or  $F \circ \gamma$  admet un extremum en 0. Ainsi, étant une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ , sa dérivée s'annule :  $\nabla F(\mathbf{x}) \cdot \gamma'(0) = \nabla F(\mathbf{x}) \cdot \mathbf{v} = 0$ .

Comme ceci est vrai pour tout vecteur orthogonal aux  $\nabla G_i(\mathbf{a})$ , il s'ensuit que  $\nabla F(\mathbf{a})$  doit être une combinaison linéaire de ces  $\nabla G_i(\mathbf{a})$ . ■

Le corollaire suivant est souvent appelé la méthode des multiplicateurs de Lagrange.

#### Corollaire 6.4.3

Sous les hypothèses du théorème 6.4.2, soit  $\mathcal{L}(\mathbf{x}, \mathbf{a}) = F(\mathbf{x}) - \mathbf{a} \cdot \mathbf{G}(\mathbf{x})$ . Si  $\mathbf{x}$  est un extremum local de  $F|_S$  alors  $\nabla \mathcal{L}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$  pour un  $\mathbf{a} \in \mathbb{R}^n$ .

DÉMONSTRATION: Il ne s'agit en fait que d'une ré-écriture : le gradient est nul si toutes les dérivées partielles (ici, par rapport aux  $x_i$  et aux  $a_j$ ) sont nulles. Les dérivées partielles par rapport aux  $x_i$  (où  $i \in \overline{m}$ ) donnent

$$\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}, \mathbf{a}) = \frac{\partial F}{\partial x_i}(\mathbf{x}) - \sum_{k \in \overline{n}} a_k \frac{\partial G_k}{\partial x_i}(\mathbf{x}) = 0,$$

ou, comme ceci est vrai  $\forall i \in \overline{m}$ , que  $\nabla F(\mathbf{x}) - \sum_{k \in \overline{n}} a_k \nabla G_k(\mathbf{x}) = \mathbf{0}$ . En d'autres termes, les conditions obtenues par le théorème 6.4.2. D'autre part, la dérivée partielle par rapport à  $a_k$  donne que

$$\frac{\partial \mathcal{L}}{\partial a_k}(\mathbf{x}, \mathbf{a}) = G_k(\mathbf{x}) = 0,$$

ou que les contraintes soient satisfaites. ■

En bref, il suffit de faire attention au fait que  $\mathcal{L} : \mathbb{R}^{m+n} \rightarrow \mathbb{R}$  et  $\nabla \mathcal{L} : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^{m+n}$  pour voir que les deux formulations sont équivalentes. Cependant, pour ne pas oublier les hypothèses dudit corollaire, la meilleure formulation est :

**Corollaire 6.4.4**

Soit  $U \subset \mathbb{R}^m$  un ouvert, soit  $F : U \rightarrow \mathbb{R}$  une fonction dérivable et  $\mathbf{G} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  une application de classe  $C^1$  telle que  $S := \mathbf{G}^{-1}(\mathbf{0}) \neq \emptyset$ . Si  $\mathbf{x}$  est un extremum local de  $F|_S$  alors

- soit  $\mathbf{x}$  n'est pas un point régulier pour  $\mathbf{G}$  (et  $\mathbf{G}(\mathbf{x}) = \mathbf{0}$ );
- soit  $\exists \mathbf{a} \in \mathbb{R}^n$  tel que  $\nabla \mathcal{L}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$ .

Cette formulation a superficiellement deux avantages :

- ne pas oublier les hypothèses de régularité ;
- mettre simultanément toutes les équations à satisfaire sur un pied d'égalité.

En effet, il est souvent avantageux de mélanger les équations de la contrainte et celles de la dépendance linéaire (de  $\nabla F$  sur les  $\nabla G_i$ ) pour résoudre le problème rapidement<sup>6</sup>. Le vrai avantage de la formulation lagrangienne se trouve dans la sous-section 6.4.iii.

Il y a ainsi  $m+n$  équations à résoudre pour  $m+n$  variables (les  $m$  variables  $x_i$  et les  $n$  "multiplicateurs"  $a_k$ ), soit

$$\begin{array}{llll} \frac{\partial \mathcal{L}}{\partial x_1} = 0 & \text{éq.B1} & \frac{\partial \mathcal{L}}{\partial a_1} = 0 & \text{éq.C1} \\ \frac{\partial \mathcal{L}}{\partial x_2} = 0 & \text{éq.B2} & \frac{\partial \mathcal{L}}{\partial a_2} = 0 & \text{éq.C2} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \frac{\partial \mathcal{L}}{\partial a_n} = 0 & \text{éq.Cn} \\ \frac{\partial \mathcal{L}}{\partial x_m} = 0 & \text{éq.Bm} & & \end{array}$$

Un peu de terminologie :

- la fonction  $\mathcal{L}$  est appelée le lagrangien ;
- les coefficients  $\mathbf{a}$  les multiplicateurs de Lagrange ;
- un point régulier  $\mathbf{x}$  où  $\exists \mathbf{a} \in \mathbb{R}^n$  tel que  $\nabla \mathcal{L}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$  est un point critique du lagrangien, ou point critique de  $\mathbf{F}$  restreinte/contrainte à  $S$  ;

6. Plutôt, par exemple, que de tenter d'isoler des variables d'abord dans les équations de la contrainte, puis de les insérer dans les autres.



— un point  $\mathbf{x}$  qui n'est pas régulier est un point singulier<sup>7</sup>.

La valeur  $F(\mathbf{x}) \in \mathbb{R}$  en un point critique  $\mathbf{x}$  est parfois appelée valeur critique.<sup>8</sup>

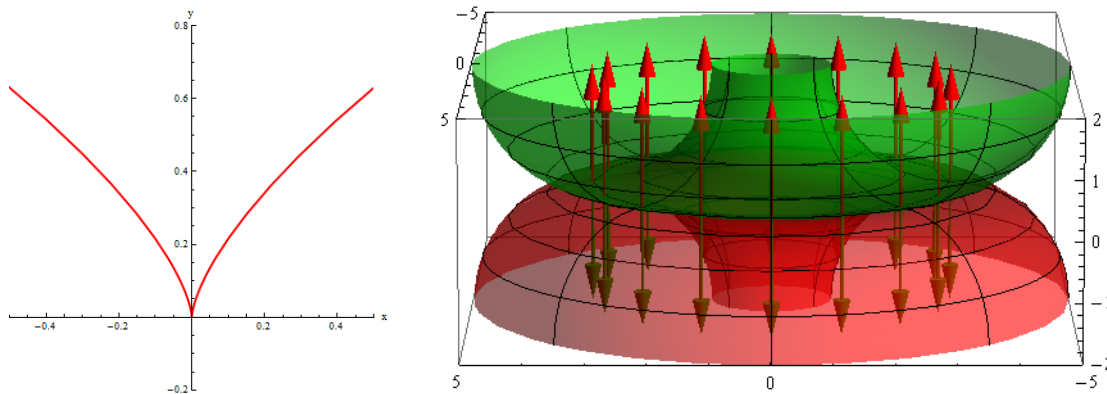
Dans ce qui suit l'acronyme “MdMdL” sera utilisé pour “Méthode des Multiplicateurs de Lagrange”, *i.e.* pour l'application du corollaire 6.4.3. Bien qu'il soit bien préférable de faire référence au corollaire 6.4.4, les habitudes veulent que la plupart des gens ne font référence qu'au corollaire 6.4.3. Le présent texte ce plie à cet us, mais non sans rappeler qu'il est TRÈS IMPORTANT de se rappeler que la MdMdL ne s'applique qu'aux points réguliers.

Voici donc un premier exemple pour montrer que cette condition de régularité n'est pas superflue.

**Exemple 6.4.5.** Soit  $F, G : \mathbb{R}^2 \rightarrow \mathbb{R}$  les fonctions définies par  $F(x, y) = y$  et  $G(x, y) = y^3 - x^2$ . Pour étudier  $F$  sous la contrainte (simple)  $S = G^{-1}(0)$ , il suffit de trouver une relation de colinéarité entre les gradients  $\nabla F(x, y) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  et  $\nabla G(x, y) = \begin{pmatrix} 2x \\ 3y^2 \end{pmatrix}$ . Or, il n'y a aucun  $a$  tel que  $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = a \begin{pmatrix} 2x \\ 3y^2 \end{pmatrix}$  et  $\begin{pmatrix} x \\ y \end{pmatrix} \in S$ .

Pour le voir, ces équations impliquent que  $3ay^2 = 1 \neq 0$ , d'où  $a \neq 0$ . Mais  $2ax = 0$ , ce qui force  $x = 0$ . Cependant, si  $x = 0$  et  $G(x, y) = 0$ , ceci implique que  $y = 0$ , une contradiction.

Pourtant,  $F|_S$  a un minimum évident en  $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Appliquer bêtement la MdMdL ne fonctionne pas car l'extremum se trouve en un point non-régulier (ou singulier), *c'ad.*  $\nabla G(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Autrement dit, ce point n'est pas un point régulier de l'ensemble de niveau ( $T_0^{vit} S = \{0\}$  tandis que  $T_0^{en} S = \mathbb{R}^2$ ) et 0 n'est pas une valeur régulière de  $G$ . Cette courbe de niveau est représentée plus bas sur la gauche. ♣



Le second est probablement plus intéressant, il montre le cas d'une contrainte qui est décrite par une application  $\mathbf{G}$  mal choisie (*c'ad.* de sorte qu'aucun point de la contrainte ne soit régulier).

**Exemple 6.4.6.** Soit  $\mathbf{G}_1 = (z - 2)^2 + (\sqrt{x^2 + y^2} - 3)^2 - 4$  la fonction (à soustraction près) du carré de la distance au cercle de rayon 3 dans le plan  $z = 2$  centré en  $(0, 0, 2)^T$  et soit  $\mathbf{G}_2 = (z + 2)^2 + (\sqrt{x^2 + y^2} - 3)^2 - 4$  la fonction (toujours à constante près) du carré de la distance au cercle de rayon 3 dans le plan  $z = -2$  centré en  $(0, 0, -2)^T$ . L'ensemble  $S = G_1^{-1}(0) \cap G_2^{-1}(0)$  est l'ensemble des points dont la distance au carré de chaque cercle est 4, soit le cercle  $x^2 + y^2 = 9$

7. Parfois, les points où  $F$  n'est pas dérivable ou  $\mathbf{G}$  n'est pas de classe  $C^1$  sont eux-aussi inclus dans cette dénomination.

8. Il arrive aussi que la valeur  $F(\mathbf{x}) \in \mathbb{R}$  en un point singulier  $\mathbf{x}$  soit appelée valeur singulière.

dans le plan  $z = 0$  (voir dessin ci-dessus à gauche). L'espace tangent de ce cercle est évidemment de dimension 1. Par contre,

$$\nabla G_1(x, y, z) = \begin{pmatrix} (1 - \frac{3}{\sqrt{x^2+y^2}})x \\ (1 - \frac{3}{\sqrt{x^2+y^2}})y \\ (z-2) \end{pmatrix} \text{ et } \nabla G_2(x, y, z) = \begin{pmatrix} (1 - \frac{3}{\sqrt{x^2+y^2}})x \\ (1 - \frac{3}{\sqrt{x^2+y^2}})y \\ (z+2) \end{pmatrix}.$$

En particulier, si  $z = 0$  et  $x^2 + y^2 = 9$ , ces gradients sont égaux à  $\begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}$  et  $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$  respectivement. Autrement dit, tout vecteur du plan  $z = 0$  leur est orthogonal, et le plan tangent en tant qu'espace de niveau est de dimension 2. Il y a alors une différence entre l'espace tangent  $T_x^{vit} S$  et  $T_x^{en} S$ ;  $\mathbf{0}$  n'est pas une valeur régulière de  $\mathbf{G}$ , et même aucun point n'est régulier pour ce choix de  $\mathbf{G}$ .

Voici ce que donne une application *erronée* de la MdMdl à l'étude de  $F(x, y, z) = x$  sur  $S = \mathbf{G}^{-1}(\mathbf{0})$ . On a  $\nabla F(x, y, z) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  n'est jamais une combinaison linéaire de  $\begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}$  et  $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ . Ce qui laisserait à penser que la fonction n'a pas d'extrema sur  $S$  (ce qui est évidemment faux, puisque  $S$  est compact). Le problème vient du fait que  $\nabla G_1$  et  $\nabla G_2$  ne sont (évidemment) pas linéairement indépendants sur  $S$ .

Par contre, il est tout à fait possible de décrire  $S$  de façon moins tordue. Par exemple, en regardant l'intersection du cylindre  $x^2 + y^2 = 9$  (l'ensemble de niveau de  $H_1(x, y, z) = x^2 + y^2 - 9$ ) et du plan  $z = 0$  (l'ensemble de niveau de  $H_2(x, y, z) = z$ ). Alors, tous les points sont réguliers, et trouver les extrema de  $F(x, y, z) = x$  se passe sans problème. ♣

Enfin, voici un exemple où toute la MdMdl se passe bien :

**Exemple 6.4.7.** Soit  $F, G : \mathbb{R}^2 \rightarrow \mathbb{R}$  les fonctions de classe  $C^1$  définies par  $F(x, y) = xe^y$  et  $G(x, y) = y - x^2$ . Tentons de trouver un minimum de  $F$  sous la contrainte  $G(x, y) = 0$ . Un calcul direct donne que  $\nabla F(x, y) = \begin{pmatrix} e^y \\ xe^y \end{pmatrix}$  et  $\nabla G(x, y) = \begin{pmatrix} 2x \\ -1 \end{pmatrix}$ . Tout d'abord,  $\nabla G(x, y)$  n'est jamais trivial ; ainsi la MdMdl s'applique.

Reste à trouver  $a \in \mathbb{R}$  tel que  $e^y = 2xa$  et  $xe^y = -a$ . Comme  $e^y > 0$ , la première de ces deux équations exclut que  $x = 0$  ou  $a = 0$ . Ainsi ces deux équations se ré-écrivent  $2xa = e^y = -\frac{a}{x}$  ce qui donne  $2x^2 = -1$ . Il n'y a donc pas de points critiques et la fonction  $F$  sous la contrainte  $G(x, y) = 0$ , en particulier,  $F$  ne possède pas d'extremum sous cette contrainte. ♣

Pour déterminer si un point critique de  $\mathcal{L}$  est un extremum local de  $F|_S$ , il n'y a pour l'instant pas d'autres méthodes que de procéder par une fonction test. Voici un corollaire important des théorèmes 6.4.2 et 2.4.9 :

#### Corollaire 6.4.8

Soit  $S \subset \mathbb{R}^m$  un ensemble compact tel que  $S = \mathbf{G}^{-1}(\mathbf{0})$  pour une application dérivable  $\mathbf{G}$  telle que les lignes de  $\mathbf{G}_*$  sont linéairement indépendantes sur  $S$ . Soit  $F$  une fonction continue sur  $S$  telle que  $F$  possède deux valeurs critiques ou singulières distinctes  $v_1 < v_2$ . Alors,  $F^{-1}(v_1)$  sont des minima absolus de  $F$  et  $F^{-1}(v_2)$  des maxima absolus de  $F$ .

DÉMONSTRATION: La démonstration en est simple : comme  $S$  est compact et  $F$  continu sur  $S$  elle doit avoir des extrema absolus. Ceux-ci ont forcément lieu en des points singuliers ou critiques. Or il n'y a que deux valeurs possibles en de tels points, ils sont tous des extrema absolus. ■

Le précédent corollaire se décline en une multitude de possibilités combinatoires, par exemple :

EXERCICE 106: Soit  $U \subset \mathbb{R}^m$  un ouvert borné tel que  $\partial U = G^{-1}(0)$  pour une fonction  $G : \mathbb{R}^m \rightarrow \mathbb{R}$  continue et de gradient non nul sur  $\partial U$ . Si  $F$  est une fonction continue sur  $U \cup \partial U$ .

1. Pourquoi  $F$  contrainte à  $\partial U$  possède-t-elle des minima absolus ? [Indice : Vous pouvez utiliser, sans démonstration, que le bord d'un ouvert est un fermé.]
2. Soit  $v$  la valeur des minima absolus de  $F$  sur  $\partial U$ . En supposant que  $F$  n'ait qu'une seule valeur critique ou singulière sur  $U$ ,  $r \in \mathbb{R}$ , montrer que si  $v > r$  alors  $F^{-1}(r)$  sont des minima absolus de  $F$ .

L'exemple suivant illustre un autre cas où l'argument de 6.4.8 s'applique indirectement.

**Exemple 6.4.9.** Soit  $\mathbf{G} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  l'application  $C^1$  définie par  $\mathbf{G}(x, y, z) = \begin{pmatrix} x^2 + y^2 - z^2 - 4 \\ x^2 + y^2 + (z-1)^2 - 9 \end{pmatrix}$  et soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y, z) = x^2 + 2y^2 + z^2$ .

Il faut d'abord regarder quand les 2 lignes de  $\mathbf{G}_*$  sont linéairement indépendantes. En fait,  $\nabla G_1(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ -2z \end{pmatrix}$  et  $\nabla G_2(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 2(z-1) \end{pmatrix}$  sont linéairement indépendants sauf en  $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  ou si  $z = \frac{1}{2}$ . Mais aucun de ces points n'appartient à  $\mathbf{G}^{-1}(\mathbf{0})$ , i.e. tous les points de  $S$  sont réguliers.

Le lagrangien est  $\mathcal{L}(x, y, z, a_1, a_2) = F(x, y, z) + a_1 G_1(x, y, z) + a_2 G_2(x, y, z)$  et son gradient s'annule lorsque

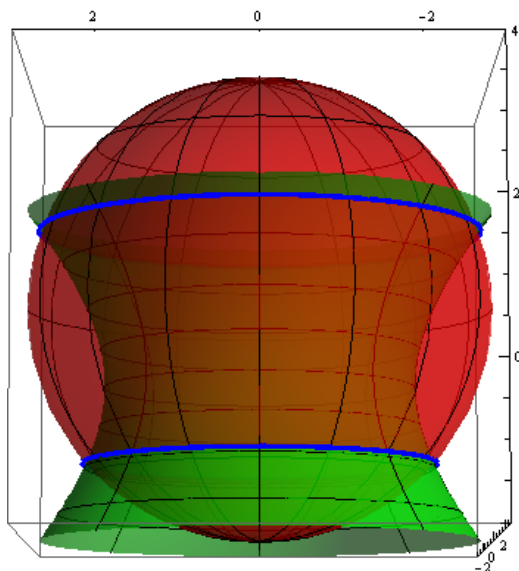
$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= 2x + 2xa_1 + 2xa_2 = 0 \Leftrightarrow 2x(1 + a_1 + a_2) = 0 & \text{éq.1} \\ \frac{\partial \mathcal{L}}{\partial y} &= 4y + 2ya_1 + 2ya_2 = 0 \Leftrightarrow 2y(2 + a_1 + a_2) = 0 & \text{éq.2} \\ \frac{\partial \mathcal{L}}{\partial z} &= 2z - 2za_1 + 2(z-1)a_2 = 0 \Leftrightarrow 2z(1 - a_1 + a_2) = 2a_2 & \text{éq.3} \\ \frac{\partial \mathcal{L}}{\partial a_1} &= x^2 + y^2 - z^2 - 4 = 0 \Leftrightarrow x^2 + y^2 = z^2 + 4 & \text{éq.4} \\ \frac{\partial \mathcal{L}}{\partial a_2} &= x^2 + y^2 + (z-1)^2 - 9 = 0 \Leftrightarrow x^2 + y^2 + (z-1)^2 = 9 & \text{éq.5} \end{aligned}$$

Tout d'abord, insérer le terme de droite de l'éq.4 dans l'éq.5 donne  $z^2 + 4 + (z-1)^2 = 9$  ou  $2z^2 - 2z - 4 = 0$  ce qui implique que  $z = \frac{1 \pm 3}{2}$ , soit  $z = 2$  ou  $z = -1$ .

De plus, l'éq.4 exclut que  $x$  et  $y$  soient tous les deux simultanément nuls. Mais si  $x$  et  $y$  sont tous les deux non-nuls, les éq.1 et éq.2 impliquent que  $1 + a_1 + a_2 = 2 + a_1 + a_2$  ce qui est impossible. Il y a ainsi deux cas :  $x = 0$  et  $y = 0$ .

Prenons d'abord  $y = 0$ . Alors, par l'éq.4  $x = \pm \sqrt{z^2 + 4}$ . Ensuite  $a_1 + a_2 = -1$  par l'éq.1. De ce fait, l'éq.3 devient  $2z(2 + 2a_2) = 2a_2$  ou  $a_2 = \frac{2z}{1-2z}$  ce qui implique  $a_1 = \frac{-1}{1-2z}$ . Voici donc les premiers points critiques de  $\mathcal{L}$  :  $(\pm\sqrt{8}, 0, 2; \frac{1}{3}, \frac{-4}{3})^T$  et  $(\pm\sqrt{5}, 0, -1; \frac{-1}{3}, \frac{-2}{3})^T$  où  $F$  prend pour valeur 12 et 6 (resp.).

Si  $x = 0$ , il apparaît par l'éq.4 que  $y = \pm \sqrt{z^2 + 4}$ . Cette fois-ci, satisfaire l'éq.2 demande à ce que  $a_1 + a_2 = -2$ . L'éq.3 devient  $2z(3 + 2a_2) = 2a_2$ , donc  $a_2 = \frac{3z}{1-2z}$  et  $a_1 = \frac{1+z}{2z-1}$ . Les points critiques de  $\mathcal{L}$  restants sont établis :  $(0, \pm\sqrt{8}, 2; 1, -2)^T$  et  $(0, \pm\sqrt{5}, -1; 0, -1)^T$  où  $F$  prend pour valeur 18 et 11 (resp.).



Il est difficile de juger quels points critiques sont des extrema (et de quel type d'extrema) ou des points d'inflexions. L'astuce ici est double. Premièrement, les contraintes impliquent que  $z = 2$  ou  $z = -1$ . Les points de  $S$  tels que  $z = 2$  sont en fait tous les  $(x, y, 2)$  tels que  $x^2 + y^2 = 8$  (voir dessin ci-contre). Il s'agit d'un ensemble fermé et borné donc compact (définition 2.2.4). Or comme  $F$  est continue, elle admet (lorsque restreinte à  $S$  et à  $z = 2$ ) un maximum et un minimum absolu sur ce compact (théorème 2.4.9). Mais attention, il ne s'agit pas de  $S$  tout entier ! Il est quand même possible d'en déduire, par le corollaire 6.4.8, les valeurs aux points critiques lorsque  $z = 2$  étant 12 et 18, que les premiers sont des minima locaux et les seconds des maxima locaux.

**[Attention ! ici il s'agit de maxima et minima absolus sur  $S \cap \{(x, y, z) \mid z = 2\}$ , sur tout  $S$  ils ne sont pour l'instant que des minima et maxima locaux.]**

*[Re-attention ! Si il y avait eu trois valeurs différentes, la plus petite aurait été associée à un [ou des] minimum[a], la plus grande à un [ou des] maximum[a], mais il aurait fallu plus de travail pour savoir si la valeur intermédiaire était effectivement un [ou des] extremum[a] ou un [ou des] points d'inflexions.]*

En faisant le même raisonnement sur  $S \cap \{(x, y, z) \mid z = -1\}$ , il apparaît que  $(0, \pm\sqrt{5}, -1)$  sont des maxima locaux (la valeur est 11) et que  $(\pm\sqrt{5}, 0, -1)$  sont des minima locaux (la valeur est 6).

Finalement, en comparant les différentes valeurs aux points critiques, il apparaît que

$\begin{pmatrix} 0 \\ \pm\sqrt{8} \\ 2 \end{pmatrix}$  sont des maxima absolus,  $\begin{pmatrix} 0 \\ \pm\sqrt{5} \\ -1 \end{pmatrix}$  sont des minima locaux,  
 $\begin{pmatrix} \pm\sqrt{8} \\ 0 \\ 2 \end{pmatrix}$  sont des maxima locaux,  $\begin{pmatrix} \pm\sqrt{5} \\ 0 \\ -1 \end{pmatrix}$  sont des minima absolus.



EXERCICE 107: Soit  $S$  un ensemble qui n'est pas nécessairement une sous-variété. Montrer que si une fonction  $F : U \rightarrow \mathbb{R}$  (où  $U$  est un ouvert contenant  $S$ ) a un maximum (resp. minimum) en  $\mathbf{x}$  alors pour tout  $\mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S$ ,  $\mathbf{v} \cdot \nabla F(\mathbf{x}) \leq 0$  (resp.  $\geq 0$ ).

Sous l'hypothèse supplémentaire que  $T_{\mathbf{x}}^{\text{vit}} S$  est symétrique (i.e.  $\mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S \Rightarrow -\mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S$ ), montrer que si  $F$  a un extremum en  $\mathbf{x}$ , alors pour tout  $\mathbf{v} \in T_{\mathbf{x}}^{\text{vit}} S$ ,  $\mathbf{v} \cdot \nabla F(\mathbf{x}) = 0$ .

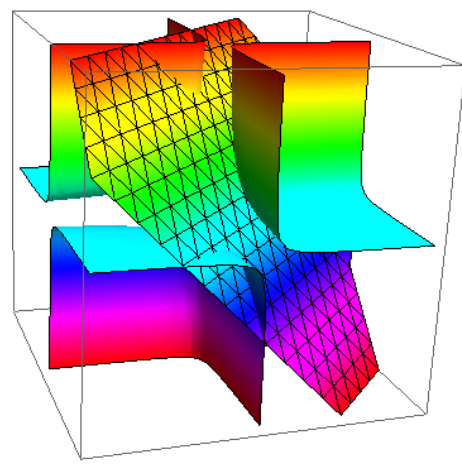
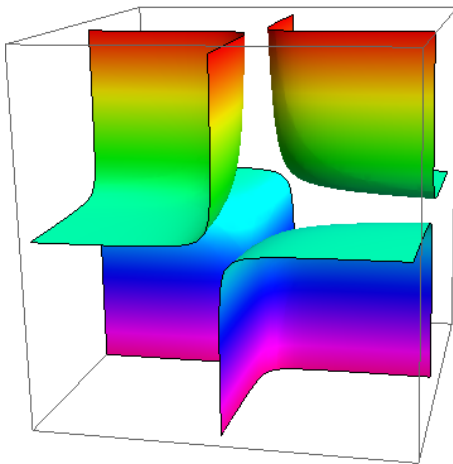
EXERCICE 108: Trouver les maxima et minima de la fonction  $F(x_1, x_2, \dots, x_m) = x_1 + x_2 + \dots + x_m = \sum_{i \in \overline{m}} x_i$  sur la sphère, i.e. sous la contrainte  $x_1^2 + x_2^2 + \dots + x_m^2 = 1$ .

EXERCICE 109: Soit  $C$  l'ensemble donné par l'intersection de la surface  $xyz = 1$  et du plan  $x + y + z = 3$ . Afin de trouver les éventuels points où l'ensemble qui en résulte a une coordonnée petite, le but ici est de chercher les extrema de  $F(x, y, z) = x$  (par symétrie, les choses se passeront pareillement

pour  $y$  et  $z$ ) restreinte à  $C$ .

- Soit  $G_1(x, y, z) = xyz - 1$  et  $G_2(x, y, z) = x + y + z - 3$ . Quels sont les points singuliers ?  
[Indice : Il y en a un d'évident. Pour les éventuels autres, utiliser que si deux vecteurs ne sont pas linéairement indépendants, tous les déterminants des mineurs de la matrice formée en les prenant pour ligne (ou colonne) sont nuls.]
- Trouver les points critiques de  $F$  sous la contrainte. (Il n'est pas nécessaire de donner le type)

[Deux dessins, le premier de l'ensemble  $xyz = 1$  ( $z$  grand en rouge), dans le second le plan  $x + y + z = 3$  (même couleur mais triangulé) est rajouté.]



EXERCICE 110: Soit  $G_1(x, y, z) = x^3 + z^3 + xy$  et  $G_2(x, y, z) = 3x + y + 3z$ . En quels points de  $C = G_1^{-1}(0) \cap G_2^{-1}(0)$  les gradients de ces deux fonctions ne sont pas linéairement indépendants ?

[Guide : Il y en a 2 (le premier provient de l'annulation d'un gradient). Pour le second, comme à l'exercice précédent exprimer cette condition par les déterminants de mineurs. Ceci permet d'exprimer  $z$  et  $y$  en fonction de  $x$ . En insérant ceci dans les contraintes, on se ramène alors à deux polynômes de degré 3 en  $\alpha = \sqrt{x}$ , puis en les combinant une équation linéaire en  $\alpha$  (ergo une solution unique). Reste à vérifier que cette valeur ( $x = \alpha^2$ ,  $y$  et  $z$  correspondants) donne bien une paire de gradient linéairement indépendants.]

#### 6.4.i Étude de fonction avec le bord du domaine.

Une autre application des multiplicateurs de Lagrange est d'étudier une fonction au bord d'un domaine. En effet, la méthode du hessien permet généralement d'étudier la fonction sur l'intérieur du domaine. Il faut ensuite tenter d'étudier la fonction sur le bord, qui étant plus petit, peut (d'habitude) se décrire comme (le bout d')un ensemble de niveau. Ainsi, étant donné  $S$  défini, par exemple, par une suite d'inégalités (strictes ou non-strictes), l'étude de  $F$  consiste à décomposer le bord de  $S$  en plusieurs morceaux  $T_i$  (chacun contenu dans un  $G_{(i)}^{-1}(0)$  pour une application  $G_{(i)}$ ), où  $i \in \bar{\ell}$  pour un  $\ell \in \mathbb{Z}_{>0}$ . Il se peut que ces morceaux aient à leur tour des bords (i.e. qu'ils soient strictement contenus dans les  $S_i$ ), ce qui complique les choses...

Finalement, il peut arriver que le bord ne soit pas contenu dans le domaine de  $F$ . Dans ce cas son étude est nécessaire seulement pour glaner quelques informations supplémentaires sur la fonction (*i.e.* est-elle bornée ?).

Voici un exemple simple :

**Exemple 6.4.10.** Soit  $S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x \geq 0 \text{ et } y > x^2 \right\} \cup \{(0,0)\}$ . Soit  $F : S \rightarrow \mathbb{R}$  la fonction de classe  $C^\infty$  définie par  $F(x,y) = y + y^2(1-x^2)$ . Pour passer à l'étude de  $F$  sur  $S$ , il faut d'abord chercher à résoudre  $\nabla F(x,y) = (-2xy^2, 1+2y(1-x^2)) = \mathbf{0}$ . La seule possibilité est  $x = 0$  et  $y = -\frac{1}{2}$ . Or ce point n'appartient pas à  $S$ . Maintenant, il faut regarder ce qui se passe au bord, *i.e.* sur  $T_1 = \{x = 0 \text{ et } y \geq 0\}$  et  $T_2 = \{y = x^2 \text{ et } y > 0\}$ .

Soit  $S_1 = G_{(1)}^{-1}(0)$  où  $G_{(1)}(x,y) = x$ . Attention :  $S_1 \supset T_1$  ! Tout d'abord  $\nabla G_{(1)}(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  ; puisqu'il ne s'annule jamais, la MdMdL s'applique. Le lagrangien associé est  $\mathcal{L}_{(1)}(x,y,a) = F(x,y) - aG_{(1)}(x,y)$ , il faut résoudre

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= -2xy^2 - ax = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= 1 + 2y(1-x^2) = 0 \\ \frac{\partial \mathcal{L}}{\partial a} &= x = 0. \end{aligned}$$

La troisième équation étant assez restrictive, il n'y a que  $x = 0$  et  $y = -\frac{1}{2}$  (de nouveau) comme solution, mais ce point n'appartient pas à  $T_1$ . Par contre,  $T_1$  a aussi un bord dans  $S_1$  : le point  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Par défaut, si une fonction contrainte à un seul point, ce point est un point critique... ici  $F(0,0) = 0$ . Pour comprendre le comportement de  $F$  au voisinage de ce point, il suffit de remarquer que, comme  $y \geq 0$  et  $x \geq 0$  dans  $S$ , si  $0 \leq x < 1$  alors  $1-x^2 > 0$  et donc  $y + y^2(1-x^2) \geq 0$  ; ainsi  $(0,0)$  est un minimum local.

Dans les régions non-bornées,  $F$  peut aussi prendre de très grandes valeurs :  $\lim_{y \rightarrow \infty} F(0,y) = +\infty$  et  $\lim_{t \rightarrow \infty} F(\sqrt{t},t) = -\infty$ . Ainsi,  $F$  ne peut pas avoir de maximum ni de minimum absolu.

En conclusion,  $F$  a un minimum local près de  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  mais ni maximum (local ou absolu) ni minimum absolu.

Pour le principe de l'exemple, voici ce qui se passe sur  $T_2$ . Soit  $S_2 = G_{(2)}^{-1}(0)$  où  $G_{(2)}(x,y) = y - x^2$  (de nouveau,  $S_2 \supset T_2$ ). Le lagrangien associé est  $\mathcal{L}_{(2)}(x,y,a) = F(x,y) - aG_{(2)}(x,y)$ , il faut résoudre

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= -2xy^2 + 2ax = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= 1 + 2y(1-x^2) - 2a = 0 \\ \frac{\partial \mathcal{L}}{\partial a} &= y - x^2 = 0. \end{aligned}$$

Comme le domaine d'intérêt est  $T_2$ , supposer que  $x > 0$  et  $y > 0$  est sans incidence. Ceci permet de simplifier le système ci-dessus en  $a = y^2$ ,  $y = x^2$ ,  $1 + 2x^2 - 4x^4 = 0$ , d'où  $x = \frac{\sqrt{1+\sqrt{5}}}{2}$ . Pour montrer qu'il s'agit d'un maximum local sur  $T_2$ , il est, entre autres, possible de regarder la fonction  $f(t) := F(\sqrt{t},t)$ , car  $\forall t > 0, (\sqrt{t},t) \in T_2$ . À remarquer que comme  $T_2 \cap S = \emptyset$  ce maximum local (sur  $T_2$ ) n'est pas important pour l'analyse en cours ; ceci permet simplement de voir le comportement de  $F$  proche de  $T_2$ . ♣

Il va sans dire que de déterminer un maximum absolu est une procédure généralement difficile (*cf.* l'exemple 6.4.9).

**Exemple 6.4.11.** Soit  $S = B_1(\mathbf{0}) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1 \right\}$  et  $F : S \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = x^2 - y^2$ . À l'intérieur de  $S$ ,  $\nabla F(x, y) = (2x, -2y)$  ne possède que  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  comme point critique. Mais comme  $F(0, 0) = 0$ ,  $F(x, 0) = x^2 > 0$  pour  $x \neq 0$  et  $F(0, y) = -y^2 < 0$  pour  $y \neq 0$ , ce point n'est en fait qu'un point d'inflexion. Comme  $S$  est compact, il doit pourtant y avoir des extrema absolus.

Si  $G(x, y) = x^2 + y^2 - 1$  alors  $\partial S = G^{-1}(0)$ . Comme  $\nabla G(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$  ne s'annule pas sur  $\partial S$ . La MdMdl s'applique donc et les solutions des équations

$$2x = 2\lambda a \text{ et } 2y = -2\lambda a$$

donneront les points critiques sur le bord. Ces équations ont pour solution (dans  $\partial S$  !)  $(x, y; a)^T = (\pm 1, 0; 1)^T$  et  $(x, y; a)^T = (0, \pm 1; -1)^T$ . Comme les valeurs associées sont 1 et -1 et que  $\partial S$  est compact, il est possible de conclure que  $(\pm 1, 0)^T$  sont des maxima absolus de  $F$  sur  $S$  et  $(0, \pm 1)^T$  des minima absolus de  $F$  sur  $S$ . ♣

Dans l'exemple précédent, il est bon de remarquer que si  $S$  avait été remplacé par son intérieur, la fonction n'aurait eu aucun extrema. Ce concept dépend donc beaucoup du domaine de définition de la fonction.

**Exemple 6.4.12.** Soit  $S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 - y^2 > 1 \right\}$  et  $F : S \rightarrow \mathbb{R}$  la fonction définie par  $F(x, y) = x^8 e^{-y^2 - x^2}$ . Son gradient est  $\nabla F(x, y) = \begin{pmatrix} x^7(8-2x^2)e^{-y^2-x^2} \\ -2x^8 y e^{-y^2-x^2} \end{pmatrix}$ ; il est trivial en  $\begin{pmatrix} \pm 2 \\ 0 \end{pmatrix}$  et en  $\begin{pmatrix} 0 \\ y \end{pmatrix}$ . Cependant, les points  $\begin{pmatrix} 0 \\ y \end{pmatrix}$  n'appartiennent pas à  $S$ .

Pour tester le type de point qu'est  $\begin{pmatrix} \pm 2 \\ 0 \end{pmatrix}$  utilisons que le développement de Taylor de  $x^8 e^{-y^2 - x^2}$  à l'ordre 2 en ces points est  $P_{2, (2,1)^T} F(h, k) = 2^8 e^{-4} - 2^9 e^{-4} h^2 - 2^8 e^{-4} k^2$ . Puisque la forme quadratique est clairement définie négative, ces points sont des maxima locaux.

$F$  pouvant très bien être définie sur tout  $\mathbb{R}$  (c'est la composition et multiplication de fonctions continues) elle s'étend à tout  $\bar{S}$  par continuité. Pour voir comment se comporte  $F$  sur  $\partial S$ , il suffit dans ce cas de voir que  $\partial S = G^{-1}(0)$  pour  $G(x, y) = x^2 - y^2 - 1$ . Alors  $\nabla G(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$  est non-nul, sauf en  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . Hors ce point n'appartient pas à  $\partial S$ . Ainsi tous les points sont réguliers, et la MdMdl s'applique. L'équation  $\nabla F(x, y) = a \nabla G(x, y)$  s'écrit  $x^7(8-2x^2)e^{-y^2-x^2} = 2\lambda a$  et  $-2x^8 y e^{-y^2-x^2} = -2\lambda a$ . Sur  $\partial S$ ,  $x \neq 0$ . Si de plus  $y \neq 0$ , il apparaît que

$$x^6(4-x^2)e^{-y^2-x^2} = a = -x^8 e^{-y^2-x^2} \Rightarrow 4x^6 = 0$$

ce qui est impossible. Donc  $y = 0$ , et comme  $x^2 - y^2 - 1 = 0$  sur  $\partial S$ , il faut que  $x = \pm 1$  (ce qui donne  $a = 3e^{-1}$ ). Ainsi,  $\begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}$  sont des points critiques de  $F$  contrainte à  $\partial S$ . En général, il peut être difficile de savoir de quel genre de point il s'agit. Dans ce cas-ci il y a une astuce : changer le signe de  $y$  ne change ni  $F$  ni  $G$ , ainsi une petite variation de la coordonnée  $y$  (et la variation associée de la coordonnée  $x$ ) aura le même effet qu'elle soit positive ou négative. Autrement dit, sous la contrainte la situation est identique à celle, en une variable, où la fonction est symétrique : le point ne peut être un point d'inflexion. Ensuite,  $\partial S$  se sépare en deux morceaux : celui de gauche ( $x < 0$ ) et celui de droite ( $x > 0$ ). Chaque composante n'a qu'un seul point critique, et, ainsi, il suffit de comparer la valeur en  $\begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}$  avec celle de n'importe quel autre point de la composante respective :

$F(\pm 1, 0) = e^{-1} > 2^8 e^{-7} = F(\pm 2, \pm \sqrt{3})$ . Ceci permet de conclure que  $\begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}$  sont des maxima locaux de  $F$  sur  $\partial S$ . Il est même possible de vérifier que lorsque  $x$  et  $y$  sont grands, les valeurs de  $F$  tendent vers 0.

Ainsi, sur  $\partial S$ ,  $F$  prend au plus la valeur de  $e^{-1}$ , et donc près du bord de  $S$  les valeurs de  $F$  restent proches de l'intervalle  $[0, e^{-1}]$ . Comme aux points  $\begin{pmatrix} \pm 2 \\ 0 \end{pmatrix}$  (dans  $S$ )  $F$  prend la valeur  $2^8 e^{-4} > e^{-1}$ , ceux-ci sont des maxima absolus de  $F$ . ♣

EXERCICE 111: Trouver les extrema de  $F(x, y, z) = xy + z^2$  sur la boule  $x^2 + y^2 + z^2 \leq 1$ . [Indice : faire une analyse sans contrainte pour classer les points à l'intérieur de la boule, puis utiliser les multiplicateurs de Lagrange pour traiter les points sur la sphère.]

EXERCICE 112: Trouver (tous) les minima et maxima de la fonction  $F : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$  définie par  $F(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^2$  sur  $\{(x, y, z, x', y', z')^T \in \mathbb{R}^6 \mid \|x = y = z, x' = 1, y' = 0, \|(x, y, z)^T - (x', y', z')^T\| \leq 10\}$ .

EXERCICE 113: L'objectif est de trouver les minima et maxima de la fonction  $F(x, y, z) = x^2 + y^2 + z^2$  restreinte à l'intersection du cône  $x^2 + y^2 = z^2$  et du plan  $x - 2z = 3$ .

- Définir deux fonctions  $G_1$  et  $G_2$  qui donne la contrainte. Dire (rapidement) pourquoi ces deux fonctions et  $F$  sont de classe  $C^1$ .
- Vérifier que les gradients de  $G_1$  et  $G_2$  sont linéairement indépendants sauf aux points de la forme  $(r, 0, 2r)^T \in \mathbb{R}^3$  où  $r \in \mathbb{R}$ .
- Montrer que la MdMdL s'applique en tout point de la contrainte, en vérifiant que les points de la forme  $(r, 0, 2r)^T$  n'appartiennent pas à la contrainte.
- Trouver les point critiques du lagrangien.
- La contrainte est un ensemble compact. Utiliser cette information (sans la démontrer !) pour déterminer quels points sont des minima et quels points sont des maxima.

Voici une méthode particulièrement tordue de retrouver un résultat classique.

EXERCICE 114: Étant donné un triangle dont la longueur des côtés est  $x$ ,  $y$  et  $z$ , son aire est

$$A(x, y, z) = \frac{1}{4} \sqrt{(x^2 + y^2 + z^2)^2 - 2(x^4 + y^4 + z^4)}.$$

Pour trouver le triangle dont l'aire est maximal à périmètre fixé revient à chercher un maximum de la fonction  $A$  sous la contrainte  $x + y + z = p$  (où  $p \in \mathbb{R}_{>0}$  est le périmètre). [Il y a évidemment d'autres méthodes pour parvenir à la conclusion de cet exercice.] Utiliser les multiplicateurs de Lagrange pour trouver quelles valeurs de  $(x, y, z)^T$  (satisfaisant cette contrainte) donnent une aire maximale. Pour ce faire, il est préférable d'étudier la fonction  $F(x, y, z) = 4A(x, y, z)^2$  et il ne faut pas oublier que seules les valeurs positives de  $x, y$  et  $z$  sont d'intérêt.

- Écrire les 4 équations que doit satisfaire les coordonnées du point critique.
- Montrer que si  $(c_1, c_2, c_3)^T$  sont les coordonnées d'un point critique, alors l'équation  $w^3 + b_1 w + b_0 = 0$  où  $b_0 = a/8$  (ici  $a$  est l'unique [coordonnée du] multiplicateur de Lagrange) et  $b_1 = -(c_1^2 + c_2^2 + c_3^2)/2$  admet pour solutions  $w = c_i$  (où  $i = 1, 2$ , ou  $3$ ).



- c. En conclure que si les trois racines sont distinctes, alors leur somme est nulle. Étant donné les contraintes, conclure que au moins deux des  $c_i$  sont égaux.
- d. Supposer (sans perdre de généralité) que  $c_1 = c_2$ . Trouver alors deux points critiques (il y a en fait trois, mais l'un a alors forcément une coordonnée négative).
- e. Montrer que (si on se restreint aux coordonnées positives) l'ensemble de la contrainte est compact. En déduire lequel des deux points critiques est le maximum et lequel le minimum.

#### 6.4.ii Inégalité arithmétique-géométrique et théorème spectral

Deux théorèmes célèbres s'obtiennent comme une application de la MdMdl. L'inégalité arithmétique-géométrique relie la moyenne (pondérée) usuelle à une moyenne dite géométrique :

##### Proposition 6.4.13

Soit  $m \in \mathbb{Z}_{>0}$ , soient  $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{R}_{>0}$  tels que  $\sum_{i \in \overline{m}} \alpha_i = 1$  et soient  $x_1, x_2, \dots, x_m \in \mathbb{R}_{>0}$ .

Alors

$$\prod_{i \in \overline{m}} x_i^{\alpha_i} \leq \sum_{i \in \overline{m}} \alpha_i x_i,$$

où l'égalité a lieu si et seulement si tous les  $x_i$  sont égaux. En particulier, si  $\alpha_i = \frac{1}{m}$ ,  $\sqrt[m]{\prod_{i \in \overline{m}} x_i} \leq \frac{1}{m} \sum_{i \in \overline{m}} x_i$

DÉMONSTRATION: Si  $m = 1$ , il faut montrer que  $x \leq x$ , ce qui est évident. Dorénavant,  $m$  sera supposé  $> 1$ . Soit  $s \in \mathbb{R}_{>0}$ , soit  $S = \{\mathbf{x} \in \mathbb{R}^m \mid \sum_{i \in \overline{m}} \alpha_i x_i = s \text{ et } \forall i \in \overline{m}, x_i \geq 0\}$  et soit  $F(\mathbf{x}) = \prod_{i \in \overline{m}} x_i^{\alpha_i}$ . Si  $G(\mathbf{x}) = -s + \sum_{i \in \overline{m}} \alpha_i x_i$ , alors  $S \subset G^{-1}(0)$ . Comme  $\nabla G(\mathbf{x}) = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$  est non-nul, la MdMdl peut être appliquée.

Pour étudier les extrema de  $F$  sous la contrainte de  $G$ , il faut poser  $\mathcal{L}(\mathbf{x}, a) = F(\mathbf{x}) - aG(\mathbf{x})$ , et chercher à résoudre

$$\begin{aligned} \forall i \in \overline{m} \quad \frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}, a) &= \frac{\alpha_i}{x_i} F(\mathbf{x}) - a\alpha_i = 0 \Leftrightarrow F(\mathbf{x}) = ax_i \\ \frac{\partial \mathcal{L}}{\partial a}(\mathbf{x}, a) &= -s + \sum_{i \in \overline{m}} \alpha_i x_i = 0 \Leftrightarrow \sum_{i \in \overline{m}} \alpha_i x_i = s. \end{aligned}$$

La première des deux équations implique soit que  $a = 0 = F(\mathbf{x})$  (et donc qu'un des  $x_i$  est nul) soit que tous les  $x_i$  sont égaux, disons à  $\bar{x}$ . Dans le deuxième cas, la seconde équation permet de conclure que  $\bar{x} = s$  (et la valeur du multiplicateur  $a$  est 1). La valeur de  $F$  en ce point critique  $(x_1, x_2, \dots, x_m)^T = (s, s, \dots, s)^T$  est  $s$ . Un point du bord de  $S$  vu comme sous-ensemble de  $G^{-1}(0)$  est un point dont une des coordonnées est nulle, ce qui coïncide avec le premier cas ;  $F$  s'annule sur  $\partial S \subset G^{-1}(0)$ . Comme  $S$  est fermé et borné (si  $m = \min \alpha_i$  alors  $S \subset B_{s/m}(\mathbf{0})$ ), il est compact. Ainsi  $(s, s, \dots, s)^T$  est un maximum absolu de  $F$  (dont la valeur est  $s$ ), ce qui implique le résultat. ■

En guise d'amuse-bouche pour le théorème spectral, la MdMdl permet aussi de trouver les axes d'une ellipse.

EXERCICE 115: Trouver les axes de l'ellipse  $3x^2 + 2xy + 3y^2 + 4z^2 = 16$  en utilisant la méthode des multiplicateurs de Lagrange.

EXERCICE 116: L'objectif de cet exercice est d'utiliser les multiplicateurs de Lagrange pour démontrer le théorème spectral sur les matrices symétriques (toutes les matrices symétriques  $n \times n$  possèdent  $n$  vecteurs propres linéairement indépendants). Soit  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  la fonction définie par  $F(\mathbf{v}) = \mathbf{v}^T Q \mathbf{v}$  et soit  $G_0 : \mathbb{R}^n \rightarrow \mathbb{R}$  définie par  $G_0(\mathbf{v}) = \mathbf{v}^T \mathbf{v}$ .

- Dire (rapidement) pourquoi  $F$  et  $G$  sont de classe  $C^\infty$ .
- Il faudra trouver les maxima de  $F$  pour la contrainte  $S^0 = G_0^{-1}(0) = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v}^T \mathbf{v} = 1\}$ . Y a-t-il des points qui ne sont pas réguliers ? Étant donné  $\mathbf{x} \in S$ , déterminer l'espace tangent  $T_{\mathbf{x}}^n S$ .
- Montrer que les points critiques du lagrangien sont des vecteurs propres de  $Q$  de valeur propre réelle.
- Dire pourquoi  $S$  est compact et en déduire l'existence d'au moins un maximum et un minimum (...distincts ?).
- Dire pourquoi il n'y a plus rien à démontrer si les deux valeurs propres sont les mêmes.
- Montrer que les vecteurs propres de valeurs propres distinctes sont orthogonaux (cet argument est un argument classique d'algèbre linéaire).
- Pour les deux vecteurs propres (de valeur propre distincte) trouvés  $\mathbf{v}_1$  et  $\mathbf{v}_2$ , ajouter une contrainte de la forme  $G_i(\mathbf{v}) = \mathbf{v}_i^T \mathbf{v} = 0$ . Montrer que l'ensemble des contraintes ne possède que des points réguliers. Quelle la dimension de  $S^{[2]} = G_0^{-1}(0) \cap G_1^{-1}(0) \cap G_2^{-1}(0)$  ?
- Pourquoi les points critiques pour cette nouvelle contrainte sont des vecteurs propres ?
- Expliquer comment itérer l'argument pour obtenir la conclusion.

### 6.4.iii Forme quadratique du lagrangien

Le lagrangien prend réellement son sens lorsqu'on tente d'étudier plus finement les points critiques, c'ad. de regarder la forme quadratique. Les choses sont plus compliquées que dans le cas sans contrainte, et le plan tangent prend alors de l'importance (c'ad. il sera non seulement utile de savoir que  $T^{\text{vit}} = T^{\text{en}} = T^{\text{par}}$ , mais aussi de trouver une base de cet espace).

Soit  $(\mathbf{x}_0, \mathbf{a}_0)$  un point critique de  $\mathcal{L}(\mathbf{x}, \mathbf{a}) = \mathbf{F}(\mathbf{x}) - \mathbf{a} \cdot \mathbf{G}(\mathbf{x})$  et supposons que  $\mathcal{L}$  soit de classe  $C^2$  dans un voisinage de ce point. En considérant le paramètre  $\mathbf{a}_0$  comme fixé, soit  $q$  la forme quadratique de  $\mathcal{L}(\cdot, \mathbf{a}_0)$  en  $\mathbf{x}_0$ . Soit  $T_{\mathbf{x}} S \subset \mathbb{R}^m$  le plan tangent à  $S = \mathbf{G}^{-1}(\mathbf{0})$  en  $\mathbf{x}$ . Alors c'est le comportement de la restriction de  $q$  aux vecteurs de l'espace tangent, qu'il faut étudier :

$$q(\mathbf{v}) = \mathbf{v}^T Q_F \mathbf{v} - \sum_{i=1}^n a_i \mathbf{v}^T Q_{G_i} \mathbf{v} \quad \text{pour } \mathbf{v} \in T_{\mathbf{x}_0}^{\text{en}} S$$

où  $Q_F$  et  $Q_{G_i}$  sont les matrices hessiennes de  $F$  et  $G_i$ .

Le calcul de cette forme quadratique se fait par le calcul de  $\{\mathbf{h}_{(1)}, \mathbf{h}_{(2)}, \dots, \mathbf{h}_{(m-n)}\}$  une base<sup>9</sup> de  $T_{\mathbf{x}} S$ . La forme quadratique du lagrangien restreinte au plan tangent est alors la forme quadratique

9. i.e. une base de l'espace vectoriel formé par les vecteurs orthogonaux à tous les  $\nabla G_j(\mathbf{x}_0)$ , où  $j \in \bar{k}$ .

de

$$H(c_1, c_2, \dots, c_{m-n}) = \mathcal{L}\left(\mathbf{x} + \sum_{j=1}^{m-n} c_j \mathbf{h}_{(j)}, \mathbf{a}\right) = F\left(\mathbf{x}_0 + \sum_{j=1}^{m-n} c_j \mathbf{h}_{(j)}\right) - \mathbf{a} \cdot \mathbf{G}\left(\mathbf{x} + \sum_{j=1}^{m-n} c_j \mathbf{h}_{(j)}\right)$$

en tant que fonction de  $c_1, c_2, \dots, c_{m-n}$ . Il est bon de noter que  $H(\mathbf{0}) = F(\mathbf{x})$  et  $\nabla H(\mathbf{0}) = \mathbf{0}$ , ainsi, il ne doit pas y avoir de terme de degré 1 dans le développement de Taylor (et le terme de degré 0 est la valeur critique). Aussi, lorsque  $F$  et  $\mathbf{G}$  sont des polynômes,  $H$  sera aussi un polynôme et sa forme quadratique est alors directement donnée par les termes de degrés deux (en les variables  $c_i$ ).

Pour résumer :

#### **Théorème 6.4.14**

Soit  $U$  un ouvert de  $\mathbb{R}^m$ ,  $F : U \rightarrow \mathbb{R}$  et  $\mathbf{G} : U \rightarrow \mathbb{R}^n$  deux applications de classe  $\mathcal{C}^2$ . Soit  $S_0 = \mathbf{G}^{-1}(\mathbf{0})$  et  $\mathbf{x} \in S_0$  un point régulier de la contrainte. De plus, soit  $\mathbf{a} \in \mathbb{R}^n$  tel que  $(\mathbf{x}, \mathbf{a})$  est un point critique de  $\mathcal{L}$ . Soit  $Q$  la forme quadratique de  $H$  comme ci-dessus.

- Si  $Q$  est définie positive (resp. négative),  $\mathbf{x}$  est un minimum (resp. maximum) local de  $F$  contrainte à  $S_0$ .
- Si  $Q$  est semi-définie positive (resp. négative) et non-triviale,  $\mathbf{x}$  n'est pas un maximum (resp. minimum) local de  $F$  contrainte à  $S_0$ .
- Si  $Q$  est non-semi définie,  $\mathbf{x}$  est un point d'inflexion  $F$  contrainte à  $S_0$ .

La démonstration est laissée en exercice :

EXERCICE 117: Avec même notations que l'énoncé du théorème.

Remarque préliminaire : le théorème 6.3.4 n'assure que l'existence de courbe  $\mathcal{C}^1$ . En fait, lorsque  $\mathbf{G}$  est  $\mathcal{C}^2$  il est possible de trouver un paramétrage  $\mathcal{C}^2$  par un raffinement du théorème d'inversion locale. Ainsi, tout point d'un voisinage de  $\mathbf{x}$  sur  $S$  est relié par une courbe  $\mathcal{C}^2$  à  $\mathbf{x}$ .

- a. Supposons que pour toute courbe  $\gamma : [-1, 1] \rightarrow S_0$  de classe  $\mathcal{C}^2$  telle que  $\gamma(0) = \mathbf{x}$ ,  $(F \circ \gamma)''(0)$  soit de même signe, montrer que le point est un maximum ou un minimum local (selon le signe).
- b. Poser  $f = F \circ \gamma(t)$  et  $\mathbf{v} = \gamma'(0)$ . Montrer que  $f''(t) = \nabla F(\mathbf{x}) \cdot \gamma''(t) + \mathbf{v}^T Q_F \mathbf{v}$  où  $Q_F$  est la matrice hessienne de  $F$ .
- c. Montrer que  $\nabla G_i(\mathbf{x}) \cdot \gamma''(t) + \mathbf{v}^T Q_{G_i} \mathbf{v} = 0$  où  $Q_{G_i}$  est la hessienne de la  $i^{\text{ème}}$  fonction coordonnée de  $\mathbf{G}$ . [Indice :  $G_i \circ \gamma \equiv 0$ .]
- d. Montrer que  $f''(t) = \mathbf{v}^T Q_F \mathbf{v} - \sum_{i=1}^n a_i \mathbf{v}^T Q_{G_i} \mathbf{v}$  pour conclure.

Tout comme à la sous-section 5.3.i “Lorsque même la dérivée seconde ne suffit plus...”, il est possible d'aller plus loin. C'est à dire, lorsque la forme quadratique est insuffisante pour conclure et que  $F$  et  $\mathbf{G}$  sont de classe  $\mathcal{C}^3$ , le polynôme homogène de degré 3 dans le développement de Taylor de  $H$  détermine le comportement. Les polynômes d'ordre 3 étant impaires, cette étape donne une conclusion facile...sauf si le polynôme s'annule dans certaines directions. Si la régularité de  $F$  et  $\mathbf{G}$  vont jusqu'à  $k$ , il est possible d'aller chercher les termes jusqu'à l'ordre  $k$  tant et aussi longtemps que les termes d'ordre inférieurs sont inconclusifs.

**Exemple 6.4.15.** Pour trouver le point de l'ensemble (fermé mais non compact)  $z^2 + yx^2 = 1$  le plus proche de l'origine, il suffit de minimiser la fonction de distance à l'origine au carré  $F(x, y, z) = x^2 + y^2 + z^2$  sous la contrainte  $G(x, y, z) = z^2 + yx^2 - 1$ .

Il faut d'abord vérifier que  $\nabla G(x, y, z) = \begin{pmatrix} 2xy \\ x^2 \\ 2z \end{pmatrix}$  ne s'annule pas sur  $G^{-1}(0)$  (et en effet, il faudrait que  $x = z = 0$ , mais alors  $G(0, y, 0) = -1$ , c'ad. ces points ne sont pas dans  $S$ ). Le lagrangien est  $\mathcal{L}(x, y, z, a) = x^2 + y^2 + z^2 - a(z^2 + yx^2 - 1)$ , et les points critiques sont les solutions de

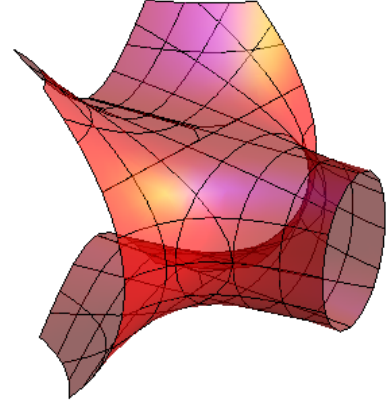
$$\frac{\partial \mathcal{L}}{\partial x} = 2x + 2axy = 0 \Leftrightarrow 2x(1 + ay) = 0 \quad \text{éq.1}$$

$$\frac{\partial \mathcal{L}}{\partial y} = 2y + ax^2 = 0 \Leftrightarrow \quad \text{éq.2}$$

$$\frac{\partial \mathcal{L}}{\partial z} = 2z + 2az = 0 \Leftrightarrow 2z(1 + a) = 0 \quad \text{éq.3}$$

$$\frac{\partial \mathcal{L}}{\partial a} = z^2 + yx^2 - 1 = 0 \Leftrightarrow z^2 + yx^2 = 1 \quad \text{éq.4}$$

De l'éq.1, il apparaît que soit  $x = 0$  soit  $ay = -1$ . Si  $x = 0$ , alors l'éq.4 donne que  $z = \pm 1$ , l'éq.2 que  $y = 0$  et l'éq.3 que  $a = -1$ . Les premiers points critiques sont donc  $(0, 0, \pm 1)$  avec  $a = -1$ .



Si maintenant  $ay = -1$ , l'éq.3 donne deux possibilités supplémentaires : soit  $z = 0$  soit  $a = -1$ . Si  $a = -1$  alors  $y = 1$ , et l'éq.2 implique que  $x = \pm\sqrt{2}$ , mais alors l'éq.4 devient  $z^2 + 2 = 1$  ce qui n'a pas de solution. Il faut donc que  $z = 0$ .

Le cas actuel est  $ay = -1$  (obtenu de l'éq.1) et  $z = 0$  (obtenu de l'éq.3). L'éq.2 donne que  $a^2x^2 = 2$ , puis l'éq.4 que  $-2 = a^3$ . D'où  $a = -2^{1/3}$  est le multiplicateur pour les points critiques  $(\pm 2^{1/6}, -2^{-1/3}, 0)^T$ .

Comme la contrainte n'est pas compacte, regarder les valeurs critiques ne peut pas suffire pour conclure. Commençons par les points  $(0, 0, \pm 1)^T$  avec  $a = -1$ . Il faut trouver la forme quadratique de  $\mathcal{L}(h, k, \pm 1 + \ell, -1)$  contrainte aux vecteurs  $(h, k, \ell)^T$  orthogonaux à  $\nabla G(0, 0, 1) = (0, 0, 2)^T$ . Il va sans dire que ces vecteurs sont les vecteurs de la forme  $(h, k, 0)^T$ . Autrement dit, il faut regarder la forme quadratique de  $\mathcal{L}(h, k, \pm 1 + 0, -1) = h^2 + k^2 + 1 - (1 + kh^2 - 1)$ . Comme ceci est déjà un polynôme, la forme quadratique (restreinte au plan tangent) est formée des termes de degré 2, soit  $h^2 + k^2$  ce qui est défini positif. Ainsi  $(0, 0, \pm 1)^T$  est un minimum local.

Maintenant, pour les points critiques  $(\pm 2^{1/6}, -2^{-1/3}, 0)^T$  avec  $a = -2^{1/3}$ , le plan tangent est donné par l'orthogonalité à  $\nabla G(\pm 2^{1/6}, -2^{-1/3}, 0) = (\mp 2^{5/6}, 2^{1/3}, 0)^T = 2^{1/3}(\mp \sqrt{2}, 1, 0)^T$ . Autrement dit, ce sont des vecteurs de la forme  $(h, \pm \sqrt{2}h, \ell)^T$ . Il ne reste qu'à trouver la forme quadratique de  $\mathcal{L}(\pm 2^{1/6} + h, -2^{-1/3} \pm \sqrt{2}h, \ell, -2^{1/3}) =$

$$(\pm 2^{1/6} + h)^2 + (-2^{-1/3} \pm \sqrt{2}h)^2 + \ell^2 - 2^{1/3}(\ell^2 + (-2^{-1/3} \pm \sqrt{2}h)(\pm 2^{1/6} + h)^2)$$

En ne gardant que les termes d'ordre 2, cette forme quadratique est :  $-(2^{1/3} - 1)\ell^2$  qui est semi-définie négative. Par conséquent, ces points ne sont pas des minima locaux (ce qui suffit pour répondre à la question initiale).

Pour ceux qui veulent en avoir le cœur net à propos du type de point critique, il faut regarder les termes d'ordres supérieurs dans le développement de  $\mathcal{L}$  en restreignant aux vecteurs pour lesquels la forme quadratique s'annule (ici, ceux où  $\ell = 0$ , i.e. de la forme  $(h, \pm\sqrt{2}h, 0)^T$ ). Dans l'équation ci-dessus, il n'y a qu'un seul terme de degré supérieur à deux, c'est  $\mp 2^{5/6}h^3$ . Celui-ci peut-être positif ou négatif selon la valeur de  $h$ , et par conséquent les points  $(\pm 2^{1/6}, -2^{-1/3}, 0)^T$  sont en réalité des points d'inflexions.

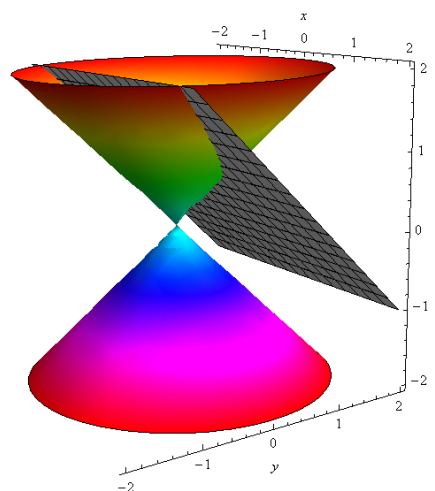
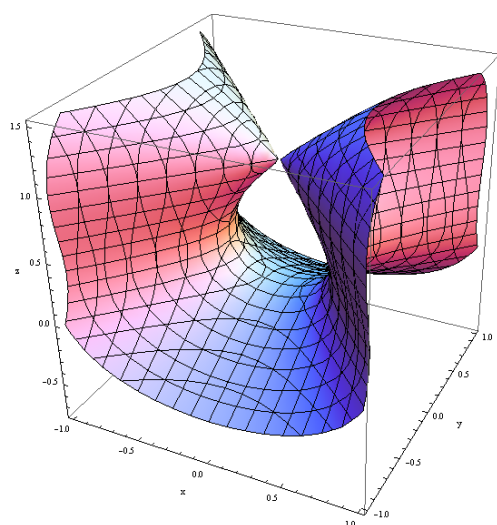


EXERCICE 118: *Le cône double est la surface (dans  $\mathbb{R}^3$ )  $S = \{\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in S \mid x_1^2 + x_2^2 = x_3^2\}$ . L'objectif est de trouver les extrema de la fonction  $F(x_1, x_2, x_3) = x_1^3 + x_2^2 + x_3^2$  (définie sur  $\mathbb{R}^3$ ) restreinte à  $S$ .*

- En quels points de  $S$  la méthode des multiplicateurs de Lagrange peut-être appliquée ?*
- Trouver les points critiques de  $F$  restreinte à  $S$ .*
- Pour un de ces deux points critiques, disons  $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ , donner une base de  $T_{\mathbf{c}}^{\text{en}} S$ .*
- Soit  $\mathbf{v}_1$  et  $\mathbf{v}_2$  les deux vecteurs de base trouvés, soit  $a \in \mathbb{R}$  le multiplicateur correspondant à  $\mathbf{c}$  et soit  $H$  le polynôme  $H(y_1, y_2) = \mathcal{L}(\mathbf{c} + y_1\mathbf{v}_1 + y_2\mathbf{v}_2, a)$ . La forme quadratique  $Q$  de ce polynôme est le polynôme homogène de degré 2 donné par tous les termes de degré 2 de  $H$ . De quel type (définie-positive, semi-définie négative, etc...) est  $Q$  ?*
- Il reste un point où la méthode des multiplicateurs de Lagrange ne s'applique pas comment procéderiez-vous pour déterminer si il correspond à un maximum ou un minimum ?*

EXERCICE 119: *Sur la surface (dans  $\mathbb{R}^3$ )  $S = \{\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in S \mid x_1^2 - x_2^2 = x_3(x_3 - 1)^2\}$ , on veut trouver les extrema de la fonction  $F(x_1, x_2, x_3) = x_1^2 + x_2^2$  (définie sur  $\mathbb{R}^3$ ) restreinte à  $S$ . Ci-dessous à gauche une représentation de la surface  $S$  :*

- En quels points de  $S$  la méthode des multiplicateurs de Lagrange peut-être appliquée ?*
- Trouver les points critiques de  $F$  restreinte à  $S$ .*
- Pour un des trois points critiques, disons  $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ , donner une base de  $T_{\mathbf{c}}^{\text{en}} S$ .*
- Soit  $\mathbf{v}_1$  et  $\mathbf{v}_2$  les deux vecteurs de base trouvés, soit  $a \in \mathbb{R}$  le multiplicateur correspondant à  $\mathbf{c}$  et soit  $H$  le polynôme  $H(y_1, y_2) = \mathcal{L}(\mathbf{c} + y_1\mathbf{v}_1 + y_2\mathbf{v}_2, a)$ . La forme quadratique  $Q$  de ce polynôme est le polynôme homogène de degré 2 donné par tous les termes de degré 2 de  $H$  (c'est la forme quadratique du lagrangien). De quel type (définie-positive, semi-définie négative, etc...) est  $Q$  ?*
- Expliquer ce qui, dans la démonstration, pourrait faire défaut dans la méthode des multiplicateurs de Lagrange aux points singuliers (i.e. les points où les hypothèses ne sont pas satisfaites). Comment procéderiez-vous pour déterminer si il correspond à un maximum ou un minimum ?*

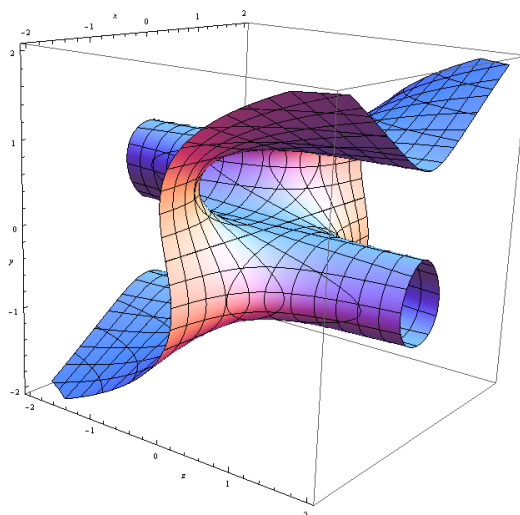
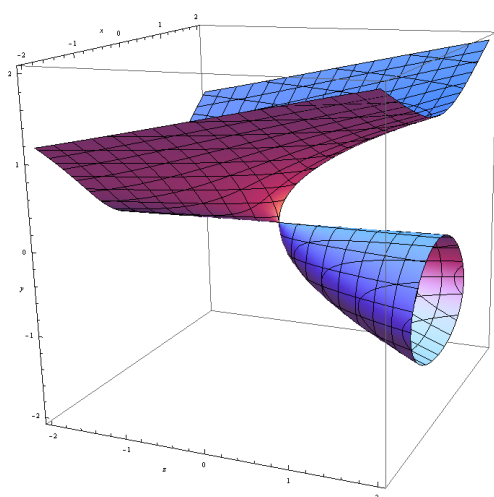


EXERCICE 120: Soient données la fonction  $F(x,y,z) = x^2 - y^2 + z$  (définie sur  $\mathbb{R}^3$ ) et les contraintes  $x^2 + y^2 = z^2$  et  $y + z = 1$  (représentées ci-dessus à droite dans  $\mathbb{R}^3$ ). Quels sont les points non-réguliers de la contrainte et les points critiques de cette fonction restreinte à la contrainte ?

Classifier les points critiques grâce à la forme quadratique du lagrangien.

EXERCICE 121: Soient données la fonction  $F(x,y,z) = y^2 - z^2$  (définie sur  $\mathbb{R}^3$ ) et la contrainte  $x^2 + yz = y^3$  (ci-dessous à gauche). Quels sont les points non-réguliers de la contrainte et les points critiques de cette fonction restreinte à la contrainte ?

Classifier les points critiques grâce à la forme quadratique du lagrangien.



EXERCICE 122: Soient données la fonction  $F(x,y,z) = y^2 - z^2$  (définie sur  $\mathbb{R}^3$ ) et la contrainte  $zx^2 + y = y^3$  (ci-dessus à droite). Quels sont les points non-réguliers de la contrainte et les points critiques de cette fonction restreinte à la contrainte ?

Classifier les points critiques grâce à la forme quadratique du lagrangien.

#### 6.4.iv Les contraintes via les paramétrages

Les théorèmes 6.3.4 et 6.3.5 montrent que la formulation en ensemble de niveau est aussi bonne que celles des paramétrages. Ainsi, il est bon de se demander pourquoi dans l'étude des fonctions, la perspective des ensembles de niveau est toujours prise. La raison tient plus à l'origine du problème qu'à une obstruction mathématique. La question de l'étude des fonctions est inspirée d'un problème : trouver le point où la fonction  $F$  (qui représente par exemple le volume d'un objet) est maximale, mais sous la contrainte (la surface de l'objet ? certains angles ou formes ?). Ces contraintes apparaissent dans la plupart des cas comme des équations ou inéquations.

D'une certaine manière, si la contrainte admet un paramétrage global, le problème est de toute façon déjà bien résolu. On peut néanmoins formuler une observation :

##### Proposition 6.4.16

*Si  $S$  est une variété qui admet un paramétrage global<sup>10</sup>  $G : U \rightarrow S$  (où  $U \subset \mathbb{R}^\ell$  est un ouvert), alors les extrema de  $F : S \rightarrow \mathbb{R}$  sont les images par  $G$  des extrema de  $F \circ G$ .*

Il est évidemment possible de prendre un recouvrement par des ouverts de  $S$ , puis de regarder des paramétrages locaux sur chaque ouvert (cela ne fait que rendre l'énonciation du résultat moins claire).

Lorsqu'un tel paramétrage existe, il y a tout les avantages à l'utiliser. La caractérisation des points critiques par les formes quadratiques est généralement plus simple que de passer par la forme quadratique du lagrangien. Cependant, trouver le paramétrage est dans la plupart des cas beaucoup plus difficile que d'appliquer la MdMdL.

## 6.A Dimension

Lorsqu'il s'agit d'espaces vectoriels, la dimension est un concept relativement facile à établir : il s'agit du plus petit nombre d'éléments nécessaires pour former une base dudit espace. Mais en parlant d'un sous-ensemble quelconque de  $\mathbb{R}^n$  est-ce vraiment clair ? Avec encore plus d'ambition, que dire d'un sous-ensemble d'un espace topologique ?

Une première et optimiste tentative pourrait être de dire qu'une bijection devrait préserver le "nombre" d'éléments et ainsi la dimension. Malheureusement, il existe une bijection entre  $[0, 1]$  et  $[0, 1]^2 = [0, 1] \times [0, 1]$  (même avec l'axiome du choix, entre  $[0, 1]$  et  $[0, 1]^\infty$ ). Cette bijection est relativement simple, la description grossière est la suivante : pour  $x \in [0, 1]$ , il y a une écriture binaire  $x = 0, b_1 b_2 b_3 \dots = \sum_{i \in \mathbb{Z}_{\geq 0}} 2^{-i} b_i$  où  $b_i \in \{0, 1\}$ . (Pour faire les choses proprement il faut s'entendre sur quelle expression choisir : pour les fractions de la forme  $\frac{n}{2^m}$  l'expression n'est pas unique, il y en a une qui se termine qu'avec des 0 et l'autre qu'avec des 1.) La fonction est alors  $f(x) = (x_{\text{pair}}, x_{\text{impair}})$  où  $x_{\text{pair}} = \sum_{i \in \mathbb{Z}_{\geq 0}} 2^{-i} x_{2i}$  et  $x_{\text{impair}} = \sum_{i \in \mathbb{Z}_{\geq 0}} 2^{-i} x_{2i+1}$ .

Une deuxième tentative serait de dire que (quand même !) si la fonction est continue, une telle aberration ne pourrait pas arriver. Notre bien-aimé feu Peano a montré que ça ne suffit pas. Il existe une courbe, *i.e.* une application  $f : [0, 1] \rightarrow [0, 1]^2$ , qui est continue et surjective.... elle s'appelle d'ailleurs la courbe de Peano.

La troisième tentative (qui est la bonne) est en fait de se dire qu'une application qui préserve la dimension doit être bijective, continue et d'inverse continue. Et enfin, personne n'est arrivé avec un contre-exemple, mais ça n'a pas aidé pour trouver une définition de la dimension. Par exemple est-ce que  $\mathbb{R} \setminus \mathbb{Q}$  est de dimension 1 ou 0 ? Il y a en fait une faune de définitions pour la dimension dont certaines coïncident selon les propriétés de l'espace topologique considéré. Soit  $X$  un ensemble et  $\mathcal{T}$  une topologie sur  $X$  (c'ad. une famille d'ouverts satisfaisant les propriétés d'une topologie).

La dimension par recouvrement est une idée de Lebesgue. Un recouvrement  $\mathcal{R}$  de  $X$  est une sous-famille de  $\mathcal{T}$  telle que  $\cup_{U \in \mathcal{R}} U = X$ . Un recouvrement  $\mathcal{R}$  est dit plus fin qu'un autre recouvrement  $\mathcal{R}'$  si  $\forall U \in \mathcal{R}, \exists U' \in \mathcal{R}'$  tel que  $U \subset U'$ ; la notation est alors  $\mathcal{R} < \mathcal{R}'$ . Le degré d'un recouvrement en  $x$  est donné par  $\deg_x \mathcal{R} = |\{U \in \mathcal{R} \mid x \in U\}|$ , c'ad. le nombre d'ouverts qui s'intersectent en  $x$ . Le degré du recouvrement  $\mathcal{R}$  tout court est  $\deg \mathcal{R} = \sup_{x \in X} \deg_x \mathcal{R}$ . Lebesgue fait la remarque suivante dans  $\mathbb{R}^n$ , pour tout recouvrement  $\mathcal{R}'$ , il est possible de trouver un recouvrement plus fin  $\mathcal{R}$  tel que  $\deg \mathcal{R} = n + 1$ . D'où l'idée  $\dim X = \sup_{\mathcal{R}' < \mathcal{T}} \inf_{\mathcal{R} < \mathcal{R}'} \deg \mathcal{R} - 1$ , le supremum étant pris sur tous les recouvrements possibles. Ce n'est par contre pas nécessairement très intuitif.

Poincaré tente (à peu de choses près) la définition suivante... qui marche relativement bien pour les sous-ensembles de  $\mathbb{R}^m$  connexes par arc. La dimension d'un ensemble est un élément de  $\mathbb{Z}_{\geq 0}$ . Une union dénombrable de points est de dimension 0. Un ensemble  $A$  connexe par arc est de dimension  $\leq n + 1$  s'il existe un ensemble  $B$  de dimension  $\leq n$  tel que  $A \setminus B$  n'est plus connexe par arc. Ça marche assez bien, mais pour les ensembles non-connexes par arc, ça devient rapidement difficile à étendre.

La dimension inductive faible (dont l'auteur ignore l'idée de qui il s'agit) est de dire que la dimension est un élément de  $\mathbb{Z}_{\geq -1}$ . L'ensemble vide a dimension  $-1$ . Un ensemble est  $A$  de dimension  $\leq n + 1$  s'il existe un ensemble  $B$  tel que  $\partial_A B$  est de dimension  $\leq n$ . Et enfin, il est possible de parler de dimension pour des sous-ensembles tordus d'espaces topologiques les plus abscons.

Un livre chaudement recommandé sur le sujet est celui de Hurewicz et Wallman, "Dimension theory", cote F10 dans la bibliothèque de l'institut de math de l'Unine. Entre autres, il y est détaillé quand  $\mathbf{F} : \mathbb{R}^m \rightarrow \mathbb{R}^n$  est continue, dans quelle mesure et dans quel sens il est possible de s'attendre à ce que  $\mathbf{F}^{-1}(\mathbf{x})$  soit de dimension  $m - n$ .

## 6.B La conjecture du Jacobien

Un problème célèbre et non-résolu en mathématiques est la conjecture du Jacobien. Étant donnée une application  $\mathbf{F} : \mathbb{C}^n \rightarrow \mathbb{C}^n$  dont toutes les fonctions coordonnées sont polynomiales, le déterminant de la matrice Jacobienne sera un polynôme. Sur les nombres complexes, ce polynôme possède forcément une racine, sauf s'il s'agit d'une constante. Si  $\mathbf{F} : \mathbb{C}^n \rightarrow \mathbb{C}^n$  est inversible, ce déterminant ne doit jamais être nul. Ainsi,

$$\mathbf{F} \text{ inversible} \Rightarrow \text{Det } \mathbf{F}_* \text{ est une constante} \neq 0.$$

D'autre part, si le Jacobien n'est jamais nul, la fonction possède un inverse localement en tout point. Un inverse local n'est cependant pas un inverse global... La conjecture du Jacobien est :

**Conjecture :** Si  $\text{Det } \mathbf{F}_*$  est une [application] constante alors  $\mathbf{F}$  possède un inverse global  $\mathbf{G}$ .



Cette affirmation n'est pas contredite même en ajoutant que la conclusion que  $\mathbf{G}$  est aussi polynomiale. Deux résultats notoires sont

- cette conjecture est vraie si les polynômes sont de degrés 2 [Wang 1980] ;
- il est suffisant de démontrer l'affirmation pour les polynômes de degré 3 pour que la conjecture soit vraie [Bass, Connell & Wright 1982].

Il va sans dire que ce problème est beaucoup plus compliqué qu'il n'en a l'air.



## Chapitre 7

# Intégrales multiples

*“Le pavé est la base de l’intégration.”*

-UN MINISTRE DE L’IMMIGRATION

De nouveau, comme la théorie de la dérivation en plusieurs variables donne en surface l’impression de ne rien faire intervenir de nouveau, celle de l’intégration se complaît dans les mêmes apparences. Il y a en réalité beaucoup de nouveautés, essentiellement reliée au fait que le bord d’un ensemble peut-être beaucoup plus irrégulier qu’en dimension 1. Pour des raisons de manque de temps, la vaste majorité des démonstrations ne sera pas présentée.

Une autre raison est que la théorie de la mesure (avec l’intégrale de Lebesgue) est beaucoup plus générale, plus simple, et plus élégante, et toute personne est fortement encouragée à se cultiver sur le sujet. L’intégrale de Riemann est une intrégrale tout à fait valable en dimension 1, mais elle est tout simplement insatisfaisante en dimension supérieure. Bien sûr, l’intérêt d’une méthode d’intégration est relative aux problèmes qui nous intéressent. Voir le texte de M. Ragolski dans la Gazette de la SMF no 137 et celui de H. Quéffelec dans la Gazette de la SMF no 132

### 7.1 Pavés

Le mot pavé est, curieusement, apparu comme la bonne dénomination pour ce domaine quelque peu grossier, mais porteur, sur lequel bâtir l’intégration en plusieurs variables quelque temps après le mois de mai 1968.

**Définition 7.1.1.** Un rectangle  $R$  dans  $\mathbb{R}^m$  est un produit cartésien de  $m$  intervalles bornés, fermés et non-vides de  $\mathbb{R}$ , *i.e.* c’est un sous-ensemble qui s’écrit  $R = \prod_{i \in \overline{m}} [a_i, b_i]$ , où  $a_i, b_i \in \mathbb{R}$  et  $b_i - a_i \geq 0$ .

Le volume d’un rectangle  $R$  est  $\text{vol}(R) = \prod_{i \in \overline{m}} (b_i - a_i) \in \mathbb{R}_{\geq 0}$ . Le volume de l’ensemble vide est (par convention) 0.

La largeur d’un rectangle  $R$  est  $\delta(R) = \max_{i \in \overline{m}} (b_i - a_i) \in \mathbb{R}_{\geq 0}$ .

Un pavé  $P$  dans  $\mathbb{R}^m$  est un rectangle de volume strictement positif. Une feuille  $P'$  dans  $\mathbb{R}^m$  est un rectangle de volume nul. ★

**Exemple 7.1.2.** Le bord d'un pavé de  $\mathbb{R}^m$  est une réunion de  $2m$  feuilles. ♣

EXERCICE 123: Montrer que l'intersection de deux rectangles est un rectangle ou l'ensemble vide. Montrer ensuite que  $\text{vol}(R_1 \cup R_2) = \text{vol}(R_1) + \text{vol}(R_2) - \text{vol}(R_1 \cap R_2)$  (si le courage fait défaut, se restreindre à  $\mathbb{R}^2$  pour avoir des calculs plus simples).

Il existe des définitions où un pavé peut être de volume nul (en particulier, même un singleton). Étant donné que les pavés sont fermés et bornés (pour  $\mathbf{x} \in P$ ,  $P \subset B_{\sqrt{n}\delta(P)}(\mathbf{x})$ ), ils sont compacts.

**Définition 7.1.3.** Un découpage d'un pavé  $P$  est une famille  $\{P_i\}_{i \in \bar{k}}$  (où  $k \in \mathbb{Z}_{>0}$ ) de pavés tels que

Dé1.  $i, j \in \bar{k}$  et  $i \neq j \Rightarrow \text{vol}(P_i \cap P_j) = 0$

Dé2.  $P = \bigcup_{i \in \bar{k}} P_i$

La largeur d'un découpage  $\mathcal{D} = \{P_i\}_{i \in \bar{k}}$  est  $\delta(\mathcal{D}) = \max_{i \in \bar{k}} \delta(P_i)$ . ★

La démarche de l'intégration de Riemann est la suivante :

1. l'intégrale d'une fonction  $F$  constante sur un pavé  $P$  (i.e.  $\exists C \in \mathbb{R}$  tel que  $\forall \mathbf{x} \in P, F(\mathbf{x}) = C$ ) est  $\int_P F(\mathbf{x}) d\text{vol} = C \text{vol}(P)$  ;
2. par linéarité, si une fonction prend une valeur constante  $C_i$  sur chaque pavé  $P_i$  d'un découpage  $\mathcal{D}$  de  $P$ , alors la formule précédente s'étend naturellement à  $\int_P F(\mathbf{x}) d\text{vol} = \sum_{i \in \bar{k}} C_i \text{vol}(P_i)$  ;
3. pour intégrer une fonction  $F$  qui n'est pas aussi simple, des découpages  $\mathcal{D}$  de plus en plus petits d'un pavé sont considérés de sorte que  $\delta(\mathcal{D})$  est de plus en plus petit. Sur chaque pavé  $P_i$ , la fonction  $F$  est remplacée par une fonction constante prenant pour valeur  $C_i = F(\mathbf{x}_{(i)})$  pour un  $\mathbf{x}_{(i)} \in P_i$ . L'intégrale de la fonction est une sorte de limite des approximations  $\sum C_i \text{vol}(P_i)$ , à la condition (restrictive !) que cette limite existe.

Pour cette dernière étape, à un découpage  $\mathcal{D} = \{P_i\}_{i \in \bar{k}}$  d'un pavé  $P$  et une fonction  $F : P \rightarrow \mathbb{R}$  bornée (i.e.  $\exists K \in \mathbb{R}_{>0}$  telle que  $F(P) \subset B_K(0)$ ) sont associés deux nombres, deux cas "extrêmes" :

$$S_1(F, \mathcal{D}) = \sum_{i \in \bar{k}} \text{vol}(P_i) \inf_{\mathbf{x} \in P_i} F(\mathbf{x}) \quad \text{et} \quad S_2(F, \mathcal{D}) = \sum_{i \in \bar{k}} \text{vol}(P_i) \sup_{\mathbf{x} \in P_i} F(\mathbf{x}),$$

les infima et suprema étant des réels (finis) car la fonction est bornée. En particulier  $S_1(\mathcal{D}, F), S_2(\mathcal{D}, F) \in [-K \text{vol}(P), K \text{vol}(P)]$ . De plus, une fonction continue sur un compact possède toujours un maximum et un minimum absolu, voir théorème 2.4.9, ainsi les suprema et infima sont alors (c'ad. lorsque la fonction est continue) réalisés par un point du pavé.

**Définition 7.1.4.** Soit  $P$  un pavé, soit  $F$  une fonction bornée sur  $P$ .  $F$  est dite intégrable sur  $P$  si

$$\lim_{\varepsilon \rightarrow 0} \inf_{\delta(\mathcal{D}) < \varepsilon} S_1(F, \mathcal{D}) = \lim_{\varepsilon \rightarrow 0} \sup_{\delta(\mathcal{D}) < \varepsilon} S_2(F, \mathcal{D}).$$

Lorsque les limites ci-dessus existent et sont égales, le réel obtenu est appelé l'intégrale (de Riemann) de  $F$  sur  $P$  et est noté  $\int_P F(\mathbf{x}) d\text{vol}$ . d'autres notations habituelles sont de remplacer  $d\text{vol}$  par  $d\mathbf{x}$  ou  $dx_1 dx_2 \cdots dx_m$  ou encore de l'enlever tout simplement :  $\int_P F = \int_P F(\mathbf{x}) d\text{vol}$ . ★

Le théorème suivant donne un critère, parfois commode, de l'intégrabilité, d'une fonction sur un pavé.

**Théorème 7.1.5**

Soit  $F$  une fonction bornée sur un pavé  $P$ . Soit  $I \in \mathbb{R}$ . Si  $\forall \varepsilon > 0, \exists \delta > 0$  tel que  $\forall \mathcal{D} = \{P_i\}_{i \in \bar{k}}$  découpage de  $P$  en pavés et  $\forall x_i \in P_i$ ,

$$\left| I - \sum_{i \in \bar{k}} f(x_i) \text{vol}(P_i) \right| < \varepsilon,$$

alors  $F$  est intégrable et son intégrale vaut  $I$ .

Voici quelques propriétés de l'intégrale de Riemann. Il est essentiellement suffisant de voir que ces propriétés sont vraies pour des fonctions qui sont constantes sur des pavés pour voir qu'elles sont vraies en général.

**Proposition 7.1.6**

Soit  $P$  un pavé,  $F, G : P \rightarrow \mathbb{R}$  deux fonctions bornées intégrables au sens de Riemann,  $a, b \in \mathbb{R}$ .

IR1. L'application  $F \mapsto \int_P F$  est linéaire, i.e.  $\int_P (aF + bG) = a \int_P F + b \int_P G$ . Le sous-ensemble des fonctions bornées qui sont intégrables est donc un espace vectoriel.

IR2.  $\text{vol}(P) \inf_{\mathbf{x} \in P} F(\mathbf{x}) \leq \int_P F \leq \text{vol}(P) \sup_{\mathbf{x} \in P} F(\mathbf{x})$ . En particulier, si  $F$  ne prend que des valeurs  $\geq 0$ , l'intégrale est  $\geq 0$ .

IR3. Si  $\forall \mathbf{x} \in P, F(\mathbf{x}) \leq G(\mathbf{x})$  alors  $\int_P F(\mathbf{x}) \leq \int_P G(\mathbf{x})$ .

IR4.  $\int_P |F| \geq \left| \int_P F \right|$ .

Il reste difficile de voir si une fonction est intégrable au sens de Riemann. Une première étape pour une meilleure compréhension est le concept d'un ensemble négligeable, c'àd. un ensemble où les valeurs de la fonction ne sont pas importantes.

**Définition 7.1.7.** Un ensemble  $N$  est dit négligeable si pour tout  $\varepsilon \in \mathbb{R}_{>0}$ ,  $\exists k \in \mathbb{Z}_{>0}$  et des pavés  $\{P_i\}_{i \in \bar{k}}$  tels que  $N \subset \cup_{i \in \bar{k}} P_i$  et la somme  $\sum_{i \in \bar{k}} \text{vol}(P_i) < \varepsilon$ .

Une fonction bornée  $F : P \rightarrow \mathbb{R}$  est dite admissible s'il existe un ensemble négligeable  $N$  telle que  $F$  est continue sur  $P \setminus N$ . ★

**Exemple 7.1.8.**

- Un point est un ensemble négligeable.
- Une feuille est un ensemble négligeable.
- Une union (ou une intersection) finie d'ensembles négligeables est négligeable.
- Le bord d'un pavé est un ensemble négligeable.
- Le sous-ensemble des fonctions bornées qui sont admissibles est un espace vectoriel, i.e. si  $F, G : P \rightarrow \mathbb{R}$  sont admissibles et  $a, b \in \mathbb{R}$  alors  $aG + bF$  est admissible.
- Si  $F, G : P \rightarrow \mathbb{R}$  sont admissibles alors  $\max(F, G)$ ,  $\min(F, G)$  et  $|F|$  sont admissibles.



**Corollaire 7.1.9**

Une fonction  $F$  admissible sur un pavé  $P$  est intégrable. De plus, si  $F$  et  $G$  sont deux fonctions admissibles qui diffèrent seulement sur un ensemble négligeable alors  $\int_P F = \int_P G$ .

Il est facile de montrer une version moins forte (où la fonction est supposée continue sur tout  $P$ ) de ce théorème en utilisant le concept de continuité uniforme (voir le lemme 7.3.8).

**Définition 7.1.10.** Une application  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  est dite uniformément continue si  $\forall \varepsilon \in \mathbb{R}_{>0}, \exists \delta \in \mathbb{R}_{>0}$  tel que  $\forall x \in S, \forall y \in B_\delta(\mathbf{x}), \mathbf{F}(\mathbf{y}) \in B_\varepsilon(\mathbf{F}(\mathbf{x}))$ . ★

Et comme en dimension 1, le corollaire suivant du théorème 2.2.5 reste valable :

**Corollaire 7.1.11**

*Une application  $F$  continue sur un compact est uniformément continue.*

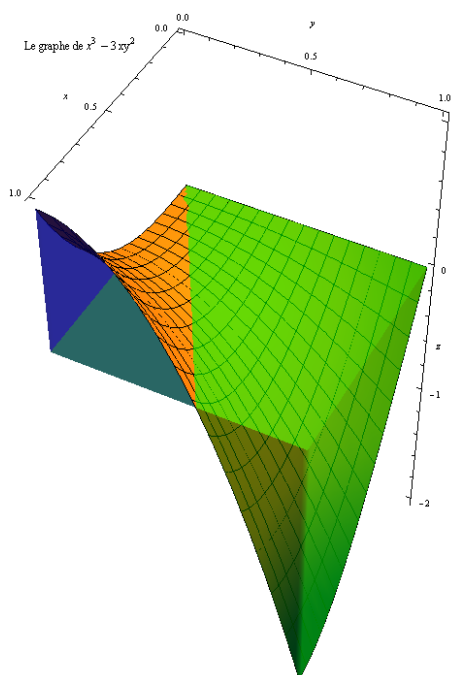
Comme les pavés sont des compacts, la continuité uniforme s'utilise en conjonction avec le théorème 7.1.5 pour montrer que toute fonction continue est intégrable.

Finalement, il est possible de montrer que la fonction  $f : [0, 1] \rightarrow \mathbb{R}$  définie par

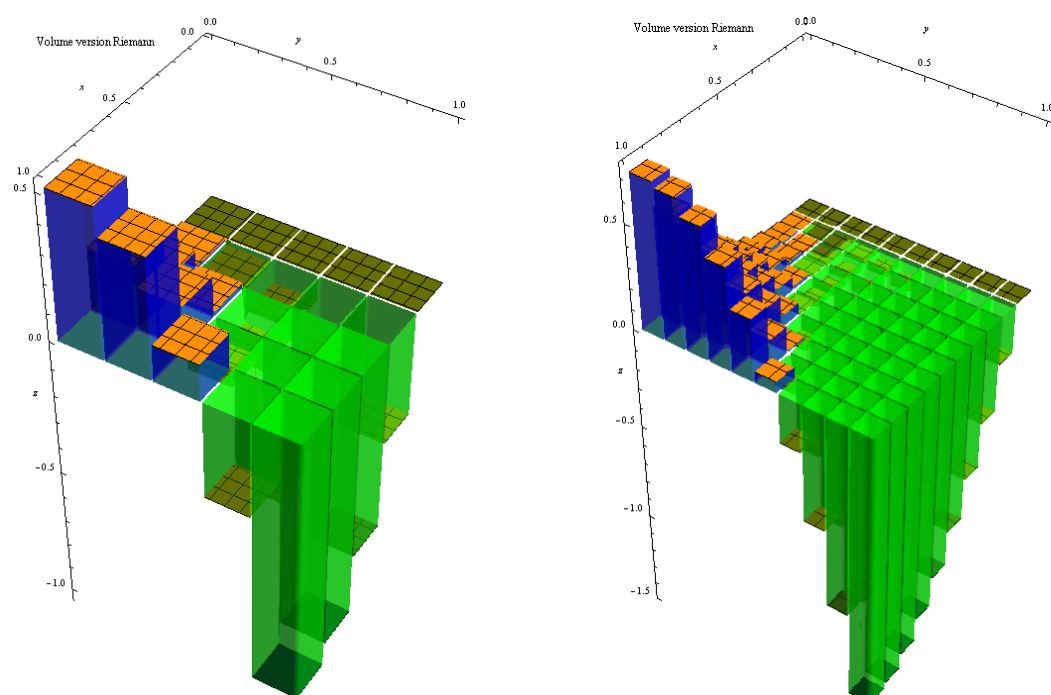
$$f(x) = \begin{cases} \frac{1}{p} & \text{si } x = \frac{q}{p} \in \mathbb{Q} \text{ où } \text{pgcd}(q, p) = 1, \\ 0 & \text{si } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

est Riemann intégrable (d'intégrale nulle).

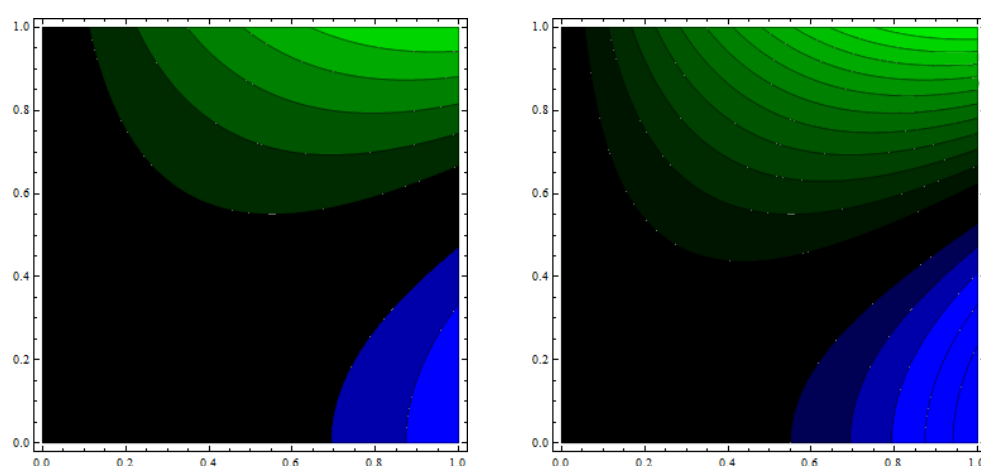
Pour mieux saisir la différence entre l'intégrale de Lebesgue et celle de Riemann, voici quelques petits dessins. Tout d'abord, la fonction à intégrer est  $F(x, y) = x^3 - 3xy^2$  disons sur le pavé  $[0, 1]^2$  (ci-dessous).



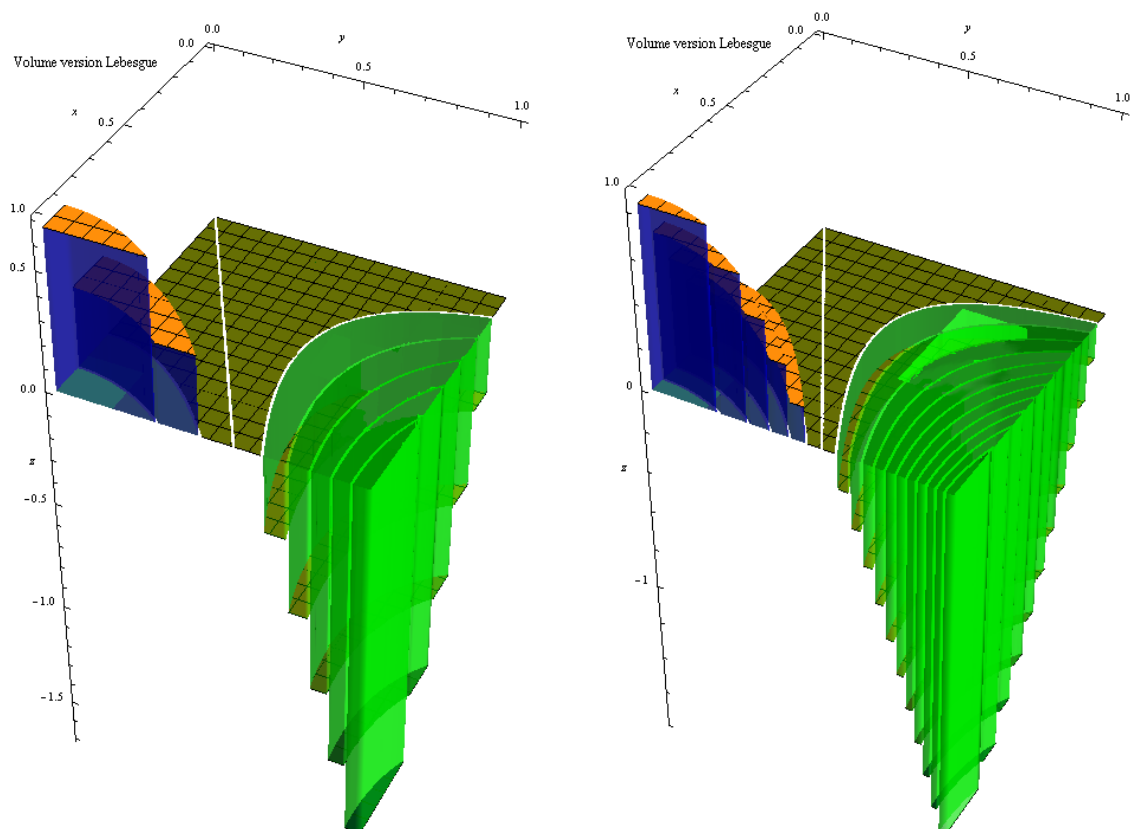
Dans l'intégrale de Riemann, il faut découper le pavé en morceaux, puis regarder la hauteur sur chaque petit pavé, ci-dessous ceci est fait pour 25 et 100 morceaux égaux (des carrés).



Dans l'intégrale de Lebesgue, la procédure est différente. Le découpage est fait sur les valeurs que prennent la fonction (ici  $F([0,1]^2) = [-2,1]$ ) ; disons que  $(I_i)_{i \in \bar{n}}$  est un découpage de  $[-2,1]$  en  $n$  morceaux (les  $I_i$  seront ici des intervalles de longueur  $3/n$ ). Tout d'abord, le domaine  $[0,1]^2$  est découpé en regardant la préimage des  $I_i$  par  $F$  :  $F^{-1}(I_i)$ . Ci-dessous ceci est fait pour 6 et 12 morceaux.



Ensuite, à chacun de ces morceaux  $F^{-1}(I_i)$  est associé le point milieu de  $I_i$  pour donner une approximation du volume.



Le problème est qu'il est difficile de donner un poids aux  $F^{-1}(I_i)$  sans déjà savoir comment parler de volume (c'est le principe de conservation des emmèlements). Cependant, ceci donne une version de l'intégrale incroyablement plus générale et plus élégante (cf. le cours de mesure).

EXERCICE 124: Montrer que la fonction  $F : [0, 1]^2 \rightarrow \mathbb{R}$  définie par

$$F(x, y) = \begin{cases} 1 & \text{si } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{Q} \times \mathbb{Q}, \\ 0 & \text{sinon,} \end{cases}$$

*n'est pas intégrable au sens de Riemann sur  $P = [0, 1]^2$ . [Indice : Montrer d'abord que  $\mathbb{Q} \times \mathbb{Q}$  est dense dans  $\mathbb{R}^2$ , i.e. que tout pavé contient un point dans  $\mathbb{Q} \times \mathbb{Q}$ . Les fonctions  $S_1(F, \mathcal{D})$  et  $S_2(F, \mathcal{D})$  s'évaluent alors pour n'importe quel découpage  $\mathcal{D}$ .]*

## 7.2 Domaines plus généraux

L'idée de base de cette section est que pour intégrer sur un sous-ensemble  $A \subset \mathbb{R}^m$  plus général une fonction (bornée)  $F : A \rightarrow \mathbb{R}$ , il est (parfois) possible d'inclure  $A$  dans un pavé  $P$ , de regarder la fonction  $F_{A \subset P} : P \rightarrow \mathbb{R}$  définie par

$$F_{A \subset P}(x) = \begin{cases} F(x) & \text{si } x \in A, \\ 0 & \text{si } x \in P \setminus A. \end{cases}$$



**Définition 7.2.1.** Un ensemble  $A$  est dit admissible s'il est borné et  $\partial A$  est négligeable. ★

Ainsi, si  $A$  est admissible,  $N \subset A$  est un ensemble négligeable, et  $F : A \rightarrow \mathbb{R}$  est continue sur  $A \setminus N$  alors, pour tout pavé  $P$  contenant  $A$ , la fonction  $F_{A \subset P} : P \rightarrow \mathbb{R}$  est admissible sur  $P$  et donc intégrable au sens de Riemann.

**Exemple 7.2.2.**

- Un pavé est admissible.
- Une intersection finie d'ensembles admissibles est admissible.
- Si  $A$  et  $B$  sont deux ensembles admissibles  $A \subset B$ , alors  $B \setminus A$  est admissible.
- Si  $A \subset \mathbb{R}^m$  et  $B \subset \mathbb{R}^n$  sont admissibles alors  $A \times B$  est admissible.



Toute boule peut être contenue dans un pavé, plus précisément,  $\forall \mathbf{x} \in \mathbb{R}^m, \forall r \in \mathbb{R}_{>0}$  alors  $B_r(\mathbf{x}) \subset \prod_{i \in \overline{m}} [x_i - r, x_i + r]$ . Ainsi, pour intégrer  $F$  sur  $A$ , il suffit de choisir un pavé  $P$  tel que  $A \subset P$  (ce pavé existe comme  $A$  est borné) et de regarder  $\int_P F_{A \subset P}$ . Il est aisé de montrer que ceci ne dépend pas du choix de pavé. Lorsque la fonction  $F_{A \subset P}$  est intégrable sur  $P$ , alors  $F$  sera dite intégrable sur  $A$  et la notation utilisée sera  $\int_A F = \int_P F_{A \subset P}$ .

Il est aussi maintenant possible d'étendre la notion de volume à n'importe quel ensemble. Soit  $\mathbb{1}_A : \mathbb{R}^m \rightarrow \mathbb{R}$  la fonction (appelée fonction caractéristique de l'ensemble  $A$ ) définie par  $\mathbb{1}_A(\mathbf{x}) = 1$  pour tout  $\mathbf{x} \in A$  et  $\mathbb{1}_A(\mathbf{x}) = 0$  si  $\mathbf{x} \notin A$ . Alors le volume de  $A$  est, pour n'importe quel pavé  $P \supset A$ ,

$$\text{vol}(A) = \int_P \mathbb{1}_A = \int_A 1.$$

Pourvu que  $F$  est intégrable sur  $A$ , toutes les propriétés de la proposition 7.1.6 valent encore pour  $P$  remplacé par  $A$ . De plus,

**Proposition 7.2.3**

Soit  $A$  et  $B$  deux ensembles admissibles,  $F$  une fonction bornée intégrable sur  $A$  et sur  $B$ , alors

1.  $F$  est intégrable sur  $A \cup B$  et  $\int_{A \cup B} F = \int_A F + \int_B F - \int_{A \cap B} F$ .
2. Si  $A$  est négligeable alors  $\int_A F = 0$ .

De la même façon que plus tôt, une fonction  $F$  sera admissible sur un ensemble admissible  $A$  s'il existe un ensemble négligeable  $N$  tel que  $F$  est continue sur  $A \setminus N$ .

**Corollaire 7.2.4**

Une fonction bornée admissible sur un ensemble admissible  $A$  est intégrable sur  $A$ .

## 7.3 Évaluation

Les deux sections précédentes ont présenté les quelques définitions et théorèmes de bases pour pouvoir parler d'intégration sur  $\mathbb{R}^m$ . Cependant, les limites sont difficiles à évaluer. Le but de cette section est de présenter les techniques qui permettent de se ramener à un cadre plus familier.

**Théorème 7.3.1 (Théorème de Fubini)**

Soit  $P$  un pavé de  $\mathbb{R}^m$  et  $Q$  un pavé de  $\mathbb{R}^n$ , soit  $F : P \times Q \rightarrow \mathbb{R}$  une fonction intégrable sur  $P \times Q$ . Supposons que,  $\forall \mathbf{x} \in P$ , la fonction  $F_{\mathbf{x}}(\mathbf{y}) = F(\mathbf{x}, \mathbf{y})$  est intégrable sur  $Q$ , alors  $\mathbf{x} \mapsto \int_Q F_{\mathbf{x}}$  est intégrable sur  $P$  et

$$\int_{P \times Q} F = \int_P \left( \int_Q F_{\mathbf{x}} \right) \quad \text{ou} \quad \int_{P \times Q} F \, d\text{vol}(\mathbf{x}, \mathbf{y}) = \int_P \left( \int_Q F(\mathbf{x}, \mathbf{y}) \, d\text{vol}(\mathbf{y}) \right) d\text{vol}(\mathbf{x}).$$

**Exemple 7.3.2.** Soit  $P = [1, 2]$  et  $Q = [-1, 2]$ .  $F(x, y) = x^3 e^y$  est continue sur  $P \times Q$ , en particulier admissible. De plus,  $\forall x \in [1, 2]$  (en fait dans  $\mathbb{R}$ ), la fonction  $y \mapsto x^3 e^y$  est admissible sur  $Q$ . Ainsi

$$\begin{aligned} \int_{P \times Q} F(x, y) \, d\text{vol}(x, y) &= \int_P \left( \int_{-1}^2 x^3 e^y \, dy \right) dx \\ &= \int_P \left( x^3 [e^y]_{-1}^2 \right) dx \\ &= (e^2 - e^{-1}) \int_1^2 x^3 \, dx \\ &= (e^2 - e^{-1}) \left( \frac{8}{3} - \frac{1}{3} \right) \end{aligned}$$

◇

♣

**Corollaire 7.3.3**

Soit  $P = \prod_{i \in \overline{m}} [a_i, b_i]$  et  $F$  une fonction continue sur  $P$  alors

$$\int_P F(\mathbf{x}) \, d\text{vol}(\mathbf{x}) = \int_{a_1}^{b_1} \left( \int_{a_2}^{b_2} \left( \cdots \left( \int_{a_m}^{b_m} F(x_1, x_2, \dots, x_m) \, dx_m \right) \cdots \right) dx_2 \right) dx_1$$

Cette deuxième forme s'appelle souvent intégrale itérée ou intégrale double (si  $m = 2$ ) ou intégrale triple (si  $m = 3$ ). Voici un petit résultat sur une forme générique d'ensembles négligeables.

**Lemme 7.3.4**

Soit  $P \subset \mathbb{R}^m$  un pavé et soit  $F : P \rightarrow \mathbb{R}$  une fonction continue, alors le graphe de  $F$  est un ensemble négligeable.

Ceci donne entre autres

**Théorème 7.3.5**

Soit  $[a, b] \subset \mathbb{R}$  un pavé (i.e.  $b - a > 0$ ). Soit  $f_1, f_2 : [a, b] \rightarrow \mathbb{R}$  deux fonctions continues telles que  $\forall x \in [a, b], f_1(x) \leq f_2(x)$ . Soit

$$A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid a \leq x \leq b \text{ et } f_1(x) \leq y \leq f_2(x) \right\}.$$

Alors  $A$  est un ensemble admissible et, si  $F$  est conforme aux hypothèses du théorème 7.3.1 de Fubini,  $\int_A F(x, y) \, d\text{vol}(x, y) = \int_a^b \left( \int_{f_1(x)}^{f_2(x)} F(x, y) \, dy \right) dx$

**Exemple 7.3.6.** Soit  $P = [-1, 1]$ ,  $f_1(x) = 0$ ,  $f_2(x) = \sqrt{1 - x^2}$  et  $A$  définie comme ci-dessus. Soit

$F(x, y) = 2x^2y$ , alors

$$\begin{aligned}\int_A F(x, y) d\text{vol}_{x, y} &= \int_{-1}^1 \left( \int_0^{\sqrt{1-x^2}} 2x^2y dy \right) dx \\ &= \int_{-1}^1 x^2 [y^2]_0^{\sqrt{1-x^2}} dx \\ &= \int_{-1}^1 x^2 (1-x^2) dx \\ &= \left[ \frac{x^3}{3} - \frac{x^5}{5} \right]_{-1}^1 \\ &= \frac{4}{15}.\end{aligned}$$

Comme d'habitude en plusieurs variables, ce même exemple peut se traiter d'une autre manière : soit  $Q = [0, 1]$ ,  $g_1(y) = -\sqrt{1-y^2}$ ,  $g_2(y) = \sqrt{1-y^2}$  et  $A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y \in [0, 1] \text{ et } g_1(y) \leq x \leq g_2(y) \right\}$ . Alors

$$\begin{aligned}\int_A F(x, y) d\text{vol}_{x, y} &= \int_0^1 \left( \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} 2x^2y dx \right) dy \\ &= \int_0^1 \frac{2}{3}y [x^3]_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dy \\ &= \int_0^1 \frac{4}{3}y (1-y^2)^{3/2} dy \\ &= \frac{4}{15} [-(1-y^2)^{5/2}/2]_0^1 \\ &= \frac{4}{15}.\end{aligned}$$

Dans ce cas-ci, les deux intégrales sont simples à évaluer, mais il arrive quelquefois qu'une des intégrales soit beaucoup plus difficile que l'autre (voire impossible). ♣

Il est possible de généraliser cette méthode en dimension supérieure. En effet, en commençant par un pavé  $[a_1, a_2] \subset \mathbb{R}$ , et  $2m$  fonctions  $f_{k,j} : I_k \rightarrow I_{k+1}$  (où  $k \in \overline{m}$  et  $j \in \overline{2}$ ) satisfaisant

- $f_{1,j} = a_j$  sont des fonctions constantes,
- $I_1 = \mathbb{R}$  et  $I_2 = [a_1, a_2]$ ,
- $I_\ell := \{ \mathbf{x} \in \mathbb{R}^\ell \mid \forall k \in \overline{\ell}, f_{k,1}(x_1, \dots, x_{k-1}) \leq x_k \leq f_{k,2}(x_1, \dots, x_{k-1}) \}$ ,
- $f_{k,1}(x_1, \dots, x_{k-1}) \leq f_{k,2}(x_1, \dots, x_{k-1})$ .

Alors l'ensemble  $A := \{ \mathbf{x} \in \mathbb{R}^m \mid \forall k \in \overline{m}, f_{k,1}(x_1 \dots x_{k-1}) \leq x_k \leq f_{k,2}(x_1, \dots, x_{k-1}) \}$  est admissible et, si  $F$  est une fonction admissible sur  $A$ , alors

$$\int_A F(\mathbf{x}) d\text{vol}(\mathbf{x}) = \int_{a_1}^{a_2} \left( \int_{f_{2,1}(x_1)}^{f_{2,2}(x_1)} \left( \int_{f_{3,1}(x_1, x_2)}^{f_{3,2}(x_1, x_2)} \left( \dots \left( \int_{f_{m,1}(x_1, \dots, x_{m-1})}^{f_{m,2}(x_1, \dots, x_{m-1})} F(x_1, x_2, \dots, x_m) dx_m \right) \dots \right) dx_3 \right) dx_2 \right) dx_1$$

Le dernier résultat, appelé principe de Cavalieri, est essentiellement un corollaire du théorème de Fubini 7.3.1. Ce principe stipule qu'il est possible d'exprimer un volume comme l'intégrale des aires de strates.

### Corollaire 7.3.7 (Principe de Cavalieri)

Soit  $A$  un sous-ensemble admissible de  $\mathbb{R}^3$ , et soit  $[a, b]$  un pavé de  $\mathbb{R}$  et  $P$  un pavé de  $\mathbb{R}^2$  tels que  $A \subset [a, b] \times P$ . Si  $\forall t \in [a, b]$ , l'ensemble  $A_t := \{ \begin{pmatrix} x \\ y \end{pmatrix} \in P \mid \begin{pmatrix} t \\ x \\ y \end{pmatrix} \in A \}$  est admissible alors

$$\text{vol}(A) = \int_a^b \text{vol}(A_t) dt.$$

EXERCICE 125: Évaluer les intégrales suivantes.

- a.  $\int_S \ln(1+xy) d\text{vol}(x, y)$  où  $S = \{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid 0 \leq x \leq y - \frac{1}{y} \text{ et } 0 < y \leq 2 \}$ .
- b.  $\int_S yz^2 e^{-xyz} d\text{vol}(x, y, z)$  où  $S = [0, 1]^3$ .

EXERCICE 126: Soit  $R \in \mathbb{R}_{>0}$ , soit  $[a, b] \subset \mathbb{R}$  un intervalle non-vidé. Utiliser les intégrales multiples pour évaluer le volume de la strate sphérique  $D := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2 \text{ et } a \leq x \leq b \right\}$ .

EXERCICE 127: Évaluer l'intégrale triple  $\int_0^1 \int_z^1 \int_0^x e^x dy dx dz$  en inversant l'ordre d'évaluation (faire un dessin du domaine, afin de le délimiter, est fortement recommandé).

EXERCICE 128: Cet exercice a pour but d'esquisser (sans trop de rigueur) une méthode d'évaluation de la primitive  $\int \frac{1}{t} \ln(1-t) dt$ .

- Pour  $x, y \in [0, 1]$ , développer la fonction  $\frac{1}{1-xy}$  en utilisant la série géométrique.
- Évaluer  $\int_S (xy)^n d\text{vol}(x, y)$  où  $n \in \mathbb{Z}_{\geq 0}$ ,  $a \in [0, 1]$  et  $S = [0, 1] \times [0, a]$ .
- En intuitionner que  $\int_S \frac{1}{1-xy} d\text{vol}(x, y) = a + \sum_{k \in \mathbb{Z}_{\geq 1}} \frac{a^{k+1}}{k^2}$ . [Il existe des critères pour montrer que l'intégrale des sommes partielles converge bien vers l'intégrale de la série.]
- Montrer d'autre part que  $\int_S \frac{1}{1-xy} d\text{vol}(x, y) = \int_0^a \frac{1}{t} \ln(1-t) dt$ .

### 7.3.i Intégrales d'applications

Il est possible de donner au sens à  $\int_A \mathbf{F}$  en faisant tout coordonnée par coordonnée. Plus précisément si  $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \dots, F_n(\mathbf{x}))^T$ , alors (pourvu que les intégrales qui seront écrites aient un sens) le vecteur  $\mathbf{v} = \int_A \mathbf{F}$  est défini par

$$\pi_i \left( \int_A \mathbf{F} \right) = \pi_i \mathbf{v} = v_i = \int_A \pi_i \mathbf{F} = \int_A F_i.$$

#### Lemme 7.3.8

Soit  $P$  un pavé et  $\mathbf{F}: P \rightarrow \mathbb{R}^n$  une application continue sur  $P$  alors

$$\left\| \int_P \mathbf{F} \right\| \leq \int_P \|\mathbf{F}\|.$$

DÉMONSTRATION: Tout d'abord, profitons de cet intermède pour montrer qu'une fonction  $G$  continue sur  $P$  est intégrable. Comme  $P$  est compact,  $G$  est uniformément continue sur  $P$ , d'où  $\forall \varepsilon \in \mathbb{R}_{>0} \exists \delta_G$  tel que  $\|\mathbf{x} - \mathbf{y}\| < \delta \Rightarrow \|G(\mathbf{x}) - G(\mathbf{y})\| < \frac{\varepsilon}{\text{vol}(P)}$ . Soit  $\mathcal{D}$  un découpage de  $P$  tel que  $\delta(\mathcal{D}) < \sqrt{n}\delta$ . Alors, pour tout  $\mathbf{x}, \mathbf{y} \in Q \in \mathcal{D}$ ,  $\|\mathbf{x} - \mathbf{y}\| < \delta$ . En particulier,  $\sup_{\mathbf{x} \in Q} G(\mathbf{x}) - \inf_{\mathbf{x} \in Q} G(\mathbf{x}) < \frac{\varepsilon}{\text{vol}(P)}$ . D'où

$$\forall \mathcal{D} \text{ tel que } \delta(\mathcal{D}) < \sqrt{n}\delta, \quad S_1(G, \mathcal{D}) - S_2(G, \mathcal{D}) < \varepsilon.$$

Ainsi, les fonctions  $F_i$  et  $\|\mathbf{F}\|$  sont intégrables (rappelons que  $\|\mathbf{F}(\mathbf{x})\|^2 = \sum_{i \in \bar{n}} |F_i(\mathbf{x})|^2$  sera continue si les  $F_i$  le sont).

Soit  $\mathcal{D}$  un découpage tel que  $\delta(\mathcal{D}) < \sqrt{n} \min(\delta_{\|\mathbf{F}\|}, \delta_{F_1}, \dots, \delta_{F_n})$ . Alors  $\forall i \in \bar{n}, \forall j \in \bar{2}, \left| \int_P F_i - \right.$

$|S_j(F_i, \mathcal{D})| < \varepsilon$  et  $|\int_P \|\mathbf{F}\| - S_j(\|\mathbf{F}\|, \mathcal{D})| < \varepsilon$ . Or

$$\begin{aligned} \left\| \left( S_2(F_1, \mathcal{D}), \dots, S_2(F_n, \mathcal{D}) \right)^T \right\| &= \left\| \left( \sum_{Q \in \mathcal{D}} \text{vol}(Q) \sup_{\mathbf{x} \in Q} F_1(\mathbf{x}), \dots, \sum_{Q \in \mathcal{D}} \text{vol}(Q) \sup_{\mathbf{x} \in Q} F_n(\mathbf{x}) \right)^T \right\| \\ &\leq \sum_{Q \in \mathcal{D}} \text{vol}(Q) \left\| \left( \sup_{\mathbf{x} \in Q} F_1(\mathbf{x}), \dots, \sup_{\mathbf{x} \in Q} F_n(\mathbf{x}) \right)^T \right\| \\ &\leq \sum_{Q \in \mathcal{D}} \text{vol}(Q) \left( \sup_{\mathbf{x} \in Q} \left\| (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))^T \right\| \right. \\ &\quad \left. + \left\| \left( \sup_{\mathbf{x}, \mathbf{y} \in Q} F_1(\mathbf{x}) - F_1(\mathbf{y}), \dots, \sup_{\mathbf{x}, \mathbf{y} \in Q} F_n(\mathbf{x}) - F_n(\mathbf{y}) \right)^T \right\| \right) \\ &< \sum_{Q \in \mathcal{D}} \text{vol}(Q) \sup_{\mathbf{x} \in Q} \|\mathbf{F}(\mathbf{x})\| + \|(\varepsilon, \dots, \varepsilon)^T\| \\ &= S_2(\|\mathbf{F}\|, \mathcal{D}) + \sqrt{n}\varepsilon \end{aligned}$$

où les passages des lignes sont justifiés par : (1<sup>ère</sup>  $\rightarrow$  2<sup>ème</sup>) inégalité du triangle (propriété No3 de la section 1.3); (2<sup>ème</sup>  $\rightarrow$  3<sup>ème</sup>) comme le sup n'est pas forcément réalisé au même point il y a une erreur possible du plus grand écart de chaque fonction coordonnée entre deux points de  $Q$ ; (3<sup>ème</sup>  $\rightarrow$  4<sup>ème</sup>) ce plus grand écart est majoré, en chaque coordonnée, par  $\frac{\varepsilon}{\text{vol}(P)}$  comme  $\delta(\mathcal{D}) < \sqrt{n}\delta_{F_j}$ .

Comme  $\left\| (S_2(F_1, \mathcal{D}), \dots, S_2(F_n, \mathcal{D}))^T - \int_P \mathbf{F} \right\| < \sqrt{n}\varepsilon$  et  $|S_2(\|\mathbf{F}\|, \mathcal{D}) - \int_P \|\mathbf{F}\|| < \varepsilon$ ,

$$\left\| \int_P \mathbf{F} \right\| < \int_P \|\mathbf{F}\| + (1 + 2\sqrt{n})\varepsilon.$$

Cependant, cette inégalité étant vraie pour tout  $\varepsilon > 0$  (et que  $n$  est fixé), ceci termine la démonstration. ■

## 7.4 Changement de variable

Lorsque les techniques précédentes ne marchent pas ou donnent lieu à des intégrales trop complexes à évaluer, il est, comme en une variable, possible et parfois préférable de faire un changement de variable. En une dimension, la condition nécessaire pour transformer une expression faisant intervenir une variable  $x$  par un changement de variable  $u = f(x)$  est que la fonction  $f$  doit être bijective, dérivable sur l'intervalle d'intégration et son inverse doit aussi être dérivable.

En plusieurs dimensions, il faut parfois changer plusieurs variables simultanément. Ce changement sera exprimé par une application  $\mathbf{F} : A \rightarrow \mathbb{R}^m$  où  $A \subset \mathbb{R}^m$  est le domaine d'intégration. (Comme la dimension à l'arrivée est identique à la dimension de départ, le terme champ de vecteur pourrait s'appliquer, cependant, il s'agit géométriquement d'un concept différent).

La formule du changement de variable donne essentiellement comment un petit intervalle (infinitésimal)  $dx$  est transformé par le changement de variable  $x = f(u)$  en un intervalle (infinitésimal) de taille  $|f'(u)|du$ . Ceci est naturel puisque la dérivée exprime justement la variation infinitésimale, et elle le fait par une approximation linéaire. Ainsi, en plusieurs variables, le point de départ pour comprendre le changement de variable sera de regarder comment le volume est transformé par une application linéaire.

### Lemme 7.4.1

Soit  $T \in \mathcal{M}_{m \times m}$  une matrice (représentant une application linéaire de  $\mathbb{R}^m$  dans  $\mathbb{R}^m$ ) et soit  $P =$

$\prod_{i \in \bar{m}} [a_i, b_i]$  un pavé de  $\mathbb{R}^m$ . Si  $TP$  est l'image de  $P$  par  $T$ , alors  $\text{vol}(TP) = |\text{Det } T| \text{vol}(P)$ .

**DÉMONSTRATION:** La preuve sera faite par réductions successives. Tout d'abord le volume étant invariant par translation, il est possible de supposer que  $P = \prod_{i \in \bar{k}} [0, b_i]$ .

Soit  $D$  la matrice diagonale dont le coefficient  $kk$  vaut  $b_k \in \mathbb{R}$  et tous les autres coefficients (non-diagonaux) valent 0. Alors, si  $Q = \prod_{i \in \bar{k}} [0, 1]$ ,  $QD = P$  et  $\text{Det } D = \text{vol}(P)$ . Ainsi, si  $\text{vol}(T'Q) = |\text{Det } T'|$  est supposée démontrée, la démonstration se termine en posant  $T' = TD$ , car le déterminant d'un produit de matrices est le produit des déterminants. Il suffit donc de montrer que la formule marche pour  $P = \prod_{i \in \bar{m}} [0, 1]$ .

Mais dans ce cas,  $PT$  est le parallélépipède engendré par les vecteurs donnés par les lignes de  $T$ , ce qui est précisément l'interprétation géométrique de  $|\text{Det } T|$ . ■

Comme la dérivée est l'approximation d'une application par une application linéaire, la formule du changement de variable (qui ne sera pas démontrée) est néanmoins intuitive :

**Théorème 7.4.2 (Changement de variable)**

Soit  $A$  un sous-ensemble admissible de  $\mathbb{R}^m$  et soit  $\mathbf{T} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  une application qui est un difféomorphisme de classe  $C^1$  sur l'intérieur de  $A$ ,  $\mathring{A}$ . Soit  $F$  une fonction admissible sur  $\mathbf{T}(A)$ , alors  $F \circ \mathbf{T}$  est admissible sur  $A$  et

$$\int_{\mathbf{T}(A)} F(\mathbf{y}) d\text{vol}(\mathbf{y}) = \int_A |\text{Det } \mathbf{T}_*(\mathbf{x})| F(\mathbf{T}(\mathbf{x})) d\text{vol}(\mathbf{x}).$$

Parmi les ingrédients de la démonstration, il faut montrer que si  $A$  est admissible et  $\mathbf{T}$  un difféomorphisme  $C^1$  alors  $\mathbf{T}(A)$  est admissible.

L'exemple suivant montre comment passer en coordonnées dites radiales, et montre comment l'utiliser pour évaluer l'intégrale  $\int_{-\infty}^{\infty} e^{-x^2} dx$  qu'il n'est pas possible d'évaluer avec les techniques d'une seule dimension.

**Exemple 7.4.3.** Soit  $\mathbf{T} : \mathbb{R}_{\geq 0} \times [-\frac{\pi}{2}, \frac{3\pi}{2}[ \rightarrow \mathbb{R}^2$  l'application donnée par  $\mathbf{T}(r, \theta) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$ . Cette application est un difféomorphisme de classe  $C^1$  sur l'intérieur ; en effet, l'application inverse est  $\mathbf{T}^{-1} : \mathbb{R}^2 \setminus \{x = 0 \text{ et } y \leq 0\} \rightarrow \mathbb{R}_{>0} \times ]-\frac{\pi}{2}, \frac{3\pi}{2}[$  définie par

$$\mathbf{T}^{-1}(x, y) = \begin{cases} (\sqrt{x^2 + y^2}, \arctg \frac{y}{x}) & \text{si } x > 0, \\ (\sqrt{x^2 + y^2}, \pi + \arctg \frac{y}{x}) & \text{si } x < 0, \\ (y, \frac{\pi}{2}) & \text{si } x = 0 \text{ (et } y > 0). \end{cases}$$

Qui plus est,  $|\text{Det } \mathbf{T}_*(r, \theta)| = r$ . Ainsi, si  $A$  est un sous-ensemble admissible de  $\mathbb{R}_{\geq 0} \times [-\frac{\pi}{2}, \frac{3\pi}{2}[$  et  $F$  une fonction admissible sur  $\mathbf{T}(A)$ ,

$$\int_{\mathbf{T}(A)} F(x, y) dx dy = \int_A F(r \cos \theta, r \sin \theta) r dr d\theta.$$

Une remarque intelligente (attribuée à Poisson<sup>1</sup>) est que pour évaluer  $\int_{\mathbb{R}} e^{-x^2} dx$ , il est plus commode d'évaluer  $\left(\int_{\mathbb{R}} e^{-x^2} dx\right)^2 = \left(\int_{\mathbb{R}} e^{-x^2} dx\right) \left(\int_{\mathbb{R}} e^{-y^2} dy\right) = \int_{\mathbb{R}^2} e^{-x^2-y^2} dx dy$  (ces manipulations sont possibles car les intégrales en question convergent absolument). Pour évaluer cette dernière intégrale, il suffit de faire le changement de variable ci-dessus :

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-x^2-y^2} dx dy &= \int_{\mathbb{R}_{\geq 0} \times [-\frac{\pi}{2}, \frac{3\pi}{2}] e^{-r^2} r dr d\theta \\ &= \int_0^\infty \left( \int_{-\pi/2}^{3\pi/2} e^{-r^2} d\theta \right) dr \\ &= 2\pi \int_0^\infty r e^{-r^2} dr \\ &= \pi [-e^{-r^2}]_0^\infty. \end{aligned}$$

D'où  $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$ . ♣

**Exemple 7.4.4.** Pour trouver l'aire du domaine borné  $A$  délimité par les paraboles  $y = x^2$ ,  $y = 2x^2$ ,  $x = y^2$  et  $x = 3y^2$ , il est aussi plus simple de procéder à un changement de variable. Soit  $\mathbf{T}$  l'application définie par  $\mathbf{T}(x, y) = \left(\frac{y}{x^2}, \frac{x}{y^2}\right)$ , alors

$$\text{Det } \mathbf{T}_*(x, y) = \begin{vmatrix} \frac{-2y}{x^3} & \frac{1}{x^2} \\ \frac{1}{y^2} & \frac{-2x}{y^3} \end{vmatrix} = \frac{3}{x^2 y^2}.$$

En particulier, comme le déterminant ne s'annule pas si  $xy \neq 0$ , l'application  $\mathbf{T}$  est, par le théorème d'inversion locale 6.2.2 un difféomorphisme  $\mathcal{C}^1$  au voisinage de tout point de  $(\mathbb{R}_{>0})^2$ . Autrement dit, pour vérifier que c'est un difféomorphisme  $\mathcal{C}^1$  sur  $(\mathbb{R}_{>0})^2$ , il ne faut que s'assurer que la fonction est bijective.

Pour se faire, le plus simple est d'inverser  $\mathbf{T}$ . En posant  $\begin{pmatrix} u \\ v \end{pmatrix} = \mathbf{T}(x, y)$ , les valeurs de  $x$  et  $y$  se retrouvent par  $x^3 = 1/u^2 v$  et  $y^3 = 1/v u^2$ .

Par surcroît,  $(\text{Det } \mathbf{T}_*)^{-1} = \text{Det}(\mathbf{T}^{-1})_*$ , et conséquemment, toujours avec  $\begin{pmatrix} u \\ v \end{pmatrix} = \mathbf{T}(x, y)$ ,

$$\text{Det } \mathbf{T}_*^{-1}(u, v) = \frac{x^2 y^2}{3} = \frac{1}{3u^2 v^2}.$$

De plus,  $\mathbf{T}(A) = [1, 2] \times [1, 3]$ , ainsi

$$\text{vol}(A) = \int_A 1 d\text{vol}(x, y) = \int_{\mathbf{T}^{-1}(\mathbf{T}(A))} 1 d\text{vol}(x, y) = \int_{\mathbf{T}(A)} |\text{Det } \mathbf{T}_*^{-1}(u, v)| d\text{vol}(u, v) = \int_1^2 \int_1^3 \frac{1}{3u^2 v^2} dv du.$$

La dernière intégrale donne facilement  $\frac{1}{9}$ . ♣

1. La méthode par les coordonnées polaires est attribuée à Poisson, plusieurs arguments permettent de deviner (plutôt que démontrer) la valeur de l'intégrale. Un argument heuristique très peu rigoureux qui permet de trouver la valeur de cette intégrale remonte à De Moivre en 1733. Pour être mené au bout (même en oubliant quelques problèmes de convergences), cet argument nécessite en plus de connaître l'approximation de Stirling, et cette dernière date de 1750.

Le premier calcul rigoureux est attribué à Laplace en 1774, et un second, du même Laplace, date de 1812. La méthode de Poisson a surtout été popularisée par Sturm (le premier écrit officiel est de 1857, elle date probablement de quelques décennies plus tôt).

Une méthode alternative de ce calcul peut-être faite en utilisant la fonction  $\Gamma$  d'Euler : cette dernière vérifie une formule dite "des compléments" qui donne essentiellement la valeur de l'intégrale. De nouveau, il a fallu attendre un peu plus longtemps avant que l'argument d'Euler puisse être fait rigoureusement.

Finalement, il y a une méthode "élémentaire" pour faire cette intégrale en restant dans le contexte d'une seule variable : voir "Evaluating the probability integral" de N. Gauthier dans la *Mathematical Gazette*, **72** (1988), mais l'origine de cette méthode est incertaine.

EXERCICE 129: Soit  $S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x > 0, y > 0, \frac{1}{2} \leq x + y \leq 1 \right\}$ .

- Montrer que  $T(u, v) = \begin{pmatrix} \frac{1}{2}(u+v) \\ \frac{1}{2}(u-v) \end{pmatrix}$  est une application bijective de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$ .
- Montrer que  $T$  est un difféomorphisme  $C^1$  de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$ , i.e. utiliser a. et montrer que  $\text{Det } T_*$  n'est jamais nul.
- Utiliser  $T$  pour calculer  $\int_S e^{\frac{x-y}{x+y}} d\text{vol}(x, y)$ .

EXERCICE 130: Soit  $D := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid (x^2 + y^2)^{1/4} \leq z \text{ et } x^2 + y^2 + z^2 \leq 2 \right\}$ . Soit  $T(r, \theta, h) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ h \end{pmatrix}$  la transformation des coordonnées cylindriques définie sur  $A = \mathbb{R}_{\geq 0} \times [0, 2\pi] \times \mathbb{R}$ . Rappelons que  $\text{Det } T_*(r, \theta, h) = r$ .

- Dire pourquoi  $T$  est un difféomorphisme de classe  $C^1$  au voisinage de tout point de l'intérieur de  $A$ .
- Décrire l'ensemble  $D'$  tel que  $T(D') = D$ .
- Évaluer le volume de  $D$  :  $\int_D 1 d\text{vol}(x, y, z)$ .

EXERCICE 131: Les coordonnées sphériques sont données par l'application  $T: \mathbb{R}_{>0} \times [0, 2\pi[ \times [0, \pi[ \rightarrow$

$$\mathbb{R}^3 \text{ définie par } T(r, \theta, \phi) = \begin{pmatrix} r \cos \theta \sin \phi \\ r \sin \theta \sin \phi \\ r \cos \phi \end{pmatrix}.$$

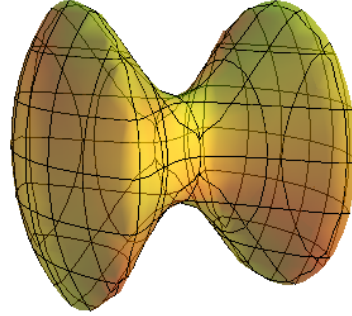
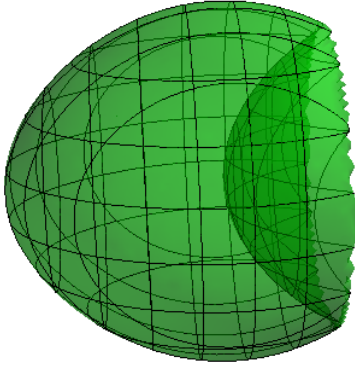
- Déterminer les formules de changement de variables pour cette application (vous pouvez supposer que cette application est bijective).
- Utiliser les intégrales multiples pour évaluer le volume de la strate sphérique  $D := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2 \text{ et } a \leq z \leq b \right\}$ .

EXERCICE 132: Évaluer le volume des régions suivantes ( $a, b, c \in \mathbb{R}$  sont des constantes).

- $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid a \leq z \leq b \text{ et } x^2 + y^2 + z^2 \leq 1 \right\}$ .
- $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \sin z \leq x^2 + y^2 \leq \cos z \right\}$ . [Indice : utiliser les coordonnées cylindriques, un changement de variable (polaire) sur  $x$  et  $y$  mais sans changer  $z$ .]
- $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 < 4 - z^2 \text{ et } x^2 + y^2 < \frac{1}{4} + z^2 \right\}$ . [Indice : idem.]
- $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid -1 \leq z/c \leq 1 \text{ et } \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} \leq 1 \right\}$  (où  $a, b, c \neq 0$ ). [Indice : utiliser une variante des coordonnées cylindriques.]

[À gauche le dessin du volume en b de l'exercice 132, à droite le volume en c du même exercice.]





EXERCICE 133: L'objectif est d'évaluer  $\int_D (x^2 - y^2)^{-1/2} d\text{vol}(x, y)$  où  $D$  est le domaine  $\{(x, y) \in \mathbb{R}^2 \mid 4 > x^2 - y^2 > 1, x > 0 \text{ et } -\frac{1}{2} < \frac{y}{x} < \frac{1}{2}\}$ .

- Soit la transformation  $T : \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \{(x, y) \mid x > |y|\}$  définie par  $T(r, t) = \begin{pmatrix} r \cosh t \\ r \sinh t \end{pmatrix}$ . Montrer que le déterminant de sa matrice de Jacobi n'est jamais nul.
- Vous pouvez utiliser sans démontrer que  $T$  est bijective. Pourquoi est-ce un difféomorphisme  $\mathcal{C}^1$  ?
- Utiliser le changement de variable pour évaluer  $\int_D (x^2 - y^2)^{-1/2} d\text{vol}(x, y)$ .

## 7.A Intégrales impropres

Il arrive souvent qu'une fonction ne soit pas bornée sur un domaine où il est nécessaire de l'intégrer, ou comme à l'exemple 7.4.3, que le domaine ne soit pas borné. Lorsque le domaine n'est pas borné, un critère relativement naturel s'applique :

### Théorème 7.A.1

Soit  $S$  un ensemble non-borné et  $\{A_k\}_{k \in \mathbb{Z}_{\geq 0}}$  une suite d'ensembles admissibles tels que  $\forall i, j \in \mathbb{Z}_{\geq 0}, i \neq j \Rightarrow A_i \cap A_j$  est négligeable et  $S = \bigcup_{k \in \mathbb{Z}_{\geq 0}} A_k$ . Soit  $F$  une fonction admissible sur tous les  $A_k$  telle que la série  $\sum_{k \in \mathbb{Z}_{\geq 0}} \int_{A_k} |F|$  converge. Alors  $\int_S F = \sum_{k \in \mathbb{Z}_{\geq 0}} \int_{A_k} F$ .

C'était en particulier le cas de la fonction  $e^{-x^2-y^2}$  sur  $\mathbb{R}^2$ .

Cependant, il n'est pas possible de relâcher un peu les hypothèses en demandant à ce qu'une fonction admissible ne soit pas nécessairement bornée en dehors d'un ensemble négligeable :

**Exemple 7.A.2.** Soit  $P = [-1, 1] \subset \mathbb{R}$  un pavé de  $\mathbb{R}$ .  $\{0\} \subset [-1, 1]$  est un ensemble négligeable. Pourtant, la fonction continue sur  $[-1, 1] \setminus \{0\}$  définie par  $f(0) = 0$  et, si  $x \neq 0$ ,  $f(x) = 1/|x|$  n'est pas intégrable sur  $[-1, 1]$ . ♣

Il n'y a essentiellement pas de technique particulière pour l'évaluation de telles intégrales. En une dimension, pour intégrer une fonction  $f$  sur un intervalle  $I$  où elle n'est pas bornée, il suffisait regarder une suite de domaines  $I_k$  plus petits tels que  $\bigcup I_k = I$  et de définir  $\int_I f = \lim_{k \rightarrow \infty} \int_{I_k} f$ .

**Définition 7.A.3.** Soit  $F$  une fonction continue sur un ensemble admissible  $A$ . Si pour toute suite croissante d'ensembles admissibles  $A_k \subset A$  tels que  $F$  est bornée sur  $A_k$ ,  $\lim_{k \rightarrow \infty} \int_{A_k} F$  existe, alors cette limite définit l'intégrale (impropre) de  $F$  sur  $A$ . ★

Soit  $A \subset \mathbb{R}^m$ , où  $m \geq 2$ , un ensemble admissible et  $F : A \rightarrow \mathbb{R}$ . Alors  $F$  est dite improprement admissible sur  $A$  si <sup>2</sup>

- $\pi_1(A)$  est un ensemble admissible de  $\mathbb{R}$ ,
- $\forall t \in \pi_1(A)$  l'ensemble  $A_t = \{\mathbf{x}' \in \mathbb{R}^{m-1} \mid \begin{pmatrix} t \\ \mathbf{x}' \end{pmatrix} \in A\}$  est admissible dans  $\mathbb{R}^{m-1}$ ,
- $\forall t \in \pi_1(A) \setminus C$ , où  $C$  est un ensemble fini, la fonction  $f_t(\mathbf{s}) = F(t, \mathbf{s})$  est admissible sur  $A_t$ ,
- l'intégrale impropre de  $g(t) = \int_{A_t} f_t(\mathbf{s}) d\text{vol}(\mathbf{s})$  converge.

**Proposition 7.A.4**

Une fonction improprement admissible (sur un ensemble admissible) admet une intégrale impropre.

**Exemple 7.A.5.** Soit  $A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x \in [0, 1] \text{ et } 0 \leq y \leq x^2 \right\}$  et soit  $F : A \rightarrow \mathbb{R}$  définie par  $F(x, y) = \frac{1}{(x+y)^2}$ . Pour  $x \neq 0$  fixé, la primitive de  $f_t(y) = F(t, y)$  existe, ainsi

$$\begin{aligned} \int_A F(x, y) &= \int_0^1 \left( \int_0^{x^2} \frac{dy}{(x+y)^2} \right) dx \\ &= \int_0^1 \left[ \frac{-1}{x+y} \right]_0^{x^2} dx \\ &= \int_0^1 \left( \frac{-1}{x+x^2} + \frac{1}{x} \right) dx \\ &= \int_0^1 \frac{1}{x+1} dx \\ &= [\ln(1+x)]_0^1 \\ &= \ln 2. \end{aligned}$$

De fait, la vérification se fait essentiellement au fur et à mesure du calcul que l'intégrale impropre existe. ♣

## 7.B Théorème de la valeur moyenne

Il s'agit d'un analogue du théorème en une dimension, à ceci près qu'il faut être plus prudent sur les domaines considérés.

**Définition 7.B.1.** Soit  $A \subset \mathbb{R}^2$  un sous-ensemble.  $A$  est dit connexe par arc si  $\forall \mathbf{x}, \mathbf{x}' \in A$ ,  $\exists \gamma : [0, 1] \rightarrow A$  une courbe paramétrée continue telle que  $\gamma(0) = \mathbf{x}$  et  $\gamma(1) = \mathbf{x}'$ . ★

**Corollaire 7.B.2**

Soit  $F : A \rightarrow \mathbb{R}$  une fonction continue sur un ensemble admissible et fermé  $A$ , alors il existe un point  $\mathbf{x}_{(0)} \in A$  tel que  $F(\mathbf{x}_{(0)}) \text{vol}(A) = \int_A F(\mathbf{x}) d\text{vol}(\mathbf{x})$ .

2. Rappel :  $\pi_1$  est la projection (canonique) sur la première coordonnée.

DÉMONSTRATION: Comme  $A$  est borné et fermé, il est compact. Ainsi  $F$  admet (au moins) un minimum (disons  $\mathbf{x}_{(1)}$ ) et un maximum (disons  $\mathbf{x}_{(2)}$ ) sur  $A$ . Ainsi

$$\text{vol}(A)F(\mathbf{x}_{(1)}) \leq \int_A F(\mathbf{x})d\text{vol}(\mathbf{x}) \leq \text{vol}(A)F(\mathbf{x}_{(2)}).$$

Soit  $\bar{F} = \text{vol}(A)^{-1} \int_A F(\mathbf{x})d\text{vol}(\mathbf{x}) \in \mathbb{R}$  la valeur moyenne, elle est comprise entre  $F(\mathbf{x}_{(1)})$  et  $F(\mathbf{x}_{(2)})$ . Puisque  $A$  est connexe par arc, il existe une courbe paramétrée  $\mathbf{G} : [0, 1] \rightarrow A$  telle que  $\mathbf{G}(0) = \mathbf{x}_{(1)}$  et  $\mathbf{G}(1) = \mathbf{x}_{(2)}$ . Alors, du théorème de la valeur moyenne pour les fonctions  $\mathbb{R} \rightarrow \mathbb{R}$  appliqué à  $F \circ \mathbf{G}$ , il existe  $t_0 \in [0, 1]$  tel que  $F \circ \mathbf{G}(t_0) = \bar{F}$ . Autrement dit  $\mathbf{x}_{(0)} = \mathbf{G}(t_0)$  est le point (de  $A$ ) qui prend la valeur moyenne. ■

Lorsque des applications à valeur dans  $\mathbb{R}^n$  ( $n > 1$ ) sont considérées, le théorème de valeur moyenne est faux :

**Exemple 7.B.3.** Soit  $\mathbf{F} : [0, 2\pi] \rightarrow \mathbb{R}^2$  l'application (une courbe paramétrée) définie par  $\mathbf{F}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ . Alors l'intégrale de chaque coordonnée est nulle, mais  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  n'est pas dans l'image de  $\mathbf{F}$ . ♣



# Chapitre 8

## Surfaces

*“La surface est essentielle pour les yeux.”*  
*“Ce qui il y a de plus profond en l’homme c’est la surface.”*  
-ANTOINE LE SAINT-EXASPÉRANT

L’objectif de ce chapitre est d’amorcer l’étude des surfaces. En bref, une surface est une sous-variété de dimension 2. Il y a cependant plusieurs perspectives possibles à prendre sur les surfaces, et certaines personnes préfèrent qualifier de surface l’image d’une “surface paramétrée”, *i.e.* l’image d’une application de  $U \subset \mathbb{R}^2$  dans  $\mathbb{R}^m$ . Plusieurs images de telles applications ne sont pas des sous-variétés (*e.g.* le cône, les surfaces qui s’intersectent elle-mêmes, etc...).

De plus, le fait d’avoir un paramétrage permet souvent de simplifier grandement l’étude de la surface. Les théorèmes 6.3.5 et 6.3.4 assurent que si une application de classe  $C^1$  est injective et que sa Jacobinne est injective en tout point, l’image sera une surface. Ils montrent aussi que toute surface admet des paramétrages locaux dans un voisinage de chaque point. Ainsi, c’est surtout le point de vue du paramétrage (ou de la surface paramétrée) qui sera pris dans ce chapitre.

Comme au chapitre 4, certaines hypothèses minimales seront faites sur les paramétrages. Ici, il sera supposé que toute surface paramétrée est une application de classe au moins  $C^1$  et que le domaine de définition  $D \subset \mathbb{R}^2$  est connexe par arc (*i.e.* toute paire de points peut être jointe par une courbe paramétrée) et ouvert<sup>1</sup>.

De la même manière que la vitesse est une caractéristique importante d’une courbe, ce sera ici la paire  $D_1\Phi$  et  $D_2\Phi$  (où  $\Phi$  est un paramétrage local<sup>2</sup>) qui sera une caractéristique importante de la surface. D’ailleurs une relecture des sections 3.7 et 6.3 est recommandée.

Afin de fixer la terminologie de ce chapitre :

- Un paramétrage (ou surface paramétrée) est une application  $C^1$  de  $D \subset \mathbb{R}^2$  (un ouvert connexe par arc) dans  $\mathbb{R}^m$
- Un paramétrage régulier  $\Phi$  est un paramétrage dont tous les points sont réguliers, *i.e.* tel que  $D_1\Phi(s, t)$  et  $D_2\Phi(s, t)$  engendrent un espace de dimension 2.

---

1. Si on veut pousser l’étude du bord des surfaces, il faut admettre des ouverts de  $\mathbb{R} \times \mathbb{R}_{\geq 0}$  (muni de sa topologie induite comme sous-ensemble de  $\mathbb{R}^2$ ).

2. Très souvent, les résultat s’appliqueront aussi lorsque  $\Phi$  n’est pas injective.

— Un paramétrage local est un paramétrage régulier injectif.

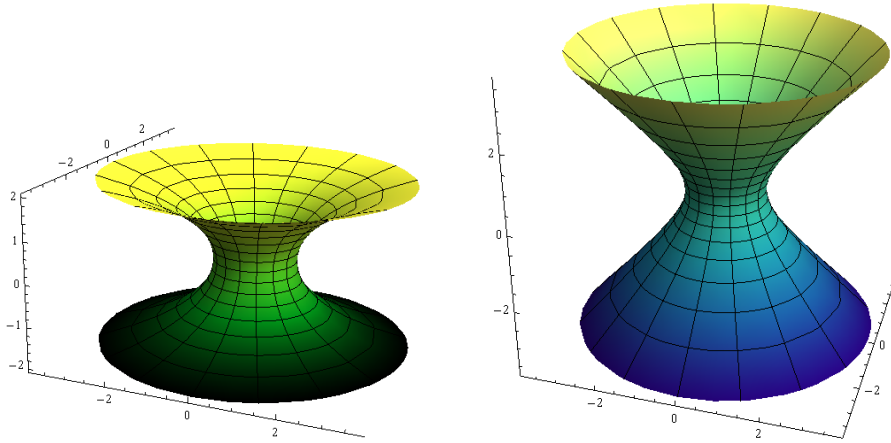
Comme pour les courbes, la notion de reparamétrage sera fréquemment utilisée.

**Définition 8.0.1.** Soit  $\Phi_{(1)} : D_1 \rightarrow \mathbb{R}^n$  et  $\Phi_{(2)} : D_2 \rightarrow \mathbb{R}^n$  deux surfaces paramétrées.  $\Phi_{(2)}$  est dit un reparamétrage  $C^k$  de  $\Phi_{(1)}$  s'il existe  $\mathbf{F} : D_2 \rightarrow D_1$  un difféomorphisme de classe  $C^k$  tel que  $\Phi_{(1)} \circ \mathbf{F} = \Phi_{(2)}$ . ★

Comme pour les courbes, deux surfaces paramétrées sont dites équivalentes si l'une est un reparamétrage  $C^1$  de l'autre (et c'est aussi une relation d'équivalence).

**Exemple 8.0.2.** Un des paramétrages du caténoïde (ci-dessous à gauche) est  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  défini par

$$\Phi(s, t) = \begin{pmatrix} \text{cht} \cos s \\ \text{cht} \sin s \\ t \end{pmatrix}.$$



Un des paramétrages de l'hyperboloïde (ci-dessus à droite) à une nappe est  $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  défini par

$$\Psi(s, t) = \begin{pmatrix} \text{cht} \cos s \\ \text{cht} \sin s \\ \text{sh}t \end{pmatrix}.$$

Ce dernier est aussi un ensemble de niveau de la fonction  $G(x, y, z) = x^2 + y^2 - z^2 : \text{Im } \Psi = G^{-1}(0)$ . Il est possible de vérifier que  $\forall (s, t), T_{\mathbf{x}}^{\text{par}} \Psi = T_{\Psi(\mathbf{x})}^{\text{en}} \text{Im } \Psi$ . ♣

EXERCICE 134: Soit  $S = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid z = x^2 - y^2 \right\}$ .

- Trouver une fonction  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  telle que  $S$  est un ensemble de niveau d'une valeur régulière de  $F$ .
- L'application  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , définie par  $\Phi(s, t) = \begin{pmatrix} s+t \\ s-t \\ 4st \end{pmatrix}$  est-elle un paramétrage global ?  
[Rappel : il faut vérifier que  $\Phi$  est bijective, qu'elle est  $C^1$ , et que les deux vecteurs donnés par ses dérivées partielles sont linéairement indépendants.]

## 8.1 Vecteur normal et orientabilité dans $\mathbb{R}^3$

Pour parler d'orientabilité d'une la surface, il est beaucoup plus intuitif de se situer d'abord dans  $\mathbb{R}^3$ . Dans  $\mathbb{R}^3$ , un plan est déterminé par un vecteur normal  $\mathbf{N}$  et un point  $\mathbf{p}$  : il est donné par tous

les points  $\mathbf{x}$  qui sont relié à  $\mathbf{p}$  par un vecteur  $\mathbf{v} = \mathbf{x} - \mathbf{p}$  tel que  $\mathbf{N} \cdot \mathbf{v} = 0$ . Lorsque le plan passe par l'origine (c'ad. c'est un sous-espace vectoriel), il est alors déterminé par le vecteur normal  $\mathbf{N}$ . Ainsi, un espace tangent de dimension 2 dans  $\mathbb{R}^3$  est déterminé par un vecteur normal.

Puisque l'espace tangent à une surface est de dimension deux, il y a toujours en un point donné la possibilité de décrire le plan tangent par un vecteur dit normal. Ceci donne aussi tout naturellement le concept d'orientabilité : est-il possible de choisir ce vecteur de sorte qu'il pointe toujours du même côté de la surface.

**Définition 8.1.1.** Une surface  $S$  est dite orientable s'il existe un champ de vecteurs normaux  $\mathbf{F} : S \rightarrow \mathbb{R}^3$  tel que

- $\mathbf{F}$  est continu ;
- $\forall \mathbf{x} \in S \subset \mathbb{R}^m, T_{\mathbf{x}}^{\text{vit}} S = \{\mathbf{v} \in \mathbb{R}^m \mid \mathbf{v} \text{ est orthogonal à } \mathbf{F}(\mathbf{x})\}$ .

La donnée de ce champ de vecteurs est une orientation de la surface. Une autre orientation est dite positive par rapport à la première si les vecteurs normaux sont positivement colinéaires, et négative s'ils sont négativement colinéaires. ★

En particulier, un champ de vecteur normal  $\mathbf{F}$  ne s'annule jamais : si c'était le cas, l'espace des vecteurs orthogonaux serait de dimension 3 et ne serait donc pas l'espace des vecteurs tangents à la surface.

L'existence d'un tel champ de vecteurs dit qu'il est possible de choisir un côté de la surface. Comme le champ de vecteurs  $\mathbf{F}$  ne s'annule pas, il est possible de le remplacer par un champ dont les vecteurs sont tous unitaires :  $\mathbf{G}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) / \|\mathbf{F}(\mathbf{x})\|$  est aussi continue car  $\|\mathbf{F}\|$  est continue et non-nulle. Il est alors clair que si la surface est orientable, il ne peut y avoir que deux orientations possibles (correspondant à un signe). Voici un résultat simple et efficace qui permet de dire qu'une surface de  $\mathbb{R}^3$  est orientable.

**Proposition 8.1.2**

Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . Soit  $S = F^{-1}(a)$  l'ensemble de niveau correspondant à la valeur régulière  $a$  (i.e. le gradient de  $F$  ne s'annule pas sur  $S$ ). Alors  $S$  est une surface orientable.

DÉMONSTRATION: Le gradient de  $F$  est orthogonal au plan tangent à la surface de niveau, cf. proposition 3.7.9, théorème 6.3.5 et théorème 6.3.4. Comme par hypothèse, il ne s'annule pas sur  $S$ , il définit une orientation de la surface. ■

**Exemple 8.1.3.** Le cylindre  $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 = 1 \right\}$  tout comme la sphère  $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1 \right\}$  ou l'hyperboloïde à une nappe  $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = 1 \right\}$  sont des surfaces orientables, puisque ce sont des surfaces de niveau d'une valeur régulière. ♣

Il est aussi important de remarquer qu'un paramétrage local (dans  $\mathbb{R}^3$ ) donne toujours un vecteur normal :

**Définition 8.1.4.** Soit  $\Phi : D \rightarrow \mathbb{R}^3$  un paramétrage local (d'une surface), alors son vecteur normal est l'application  $\mathbf{N}_{\Phi}(s, t) = D_1\Phi(s, t) \wedge D_2\Phi(s, t)$ . ★

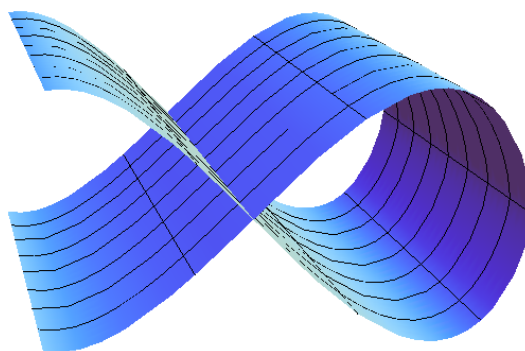
Comme il s'agit d'un paramétrage local, les deux vecteurs  $D_1\Phi$  et  $D_2\Phi$  sont linéairement indépendants. Ainsi,  $\mathbf{N}$  est non-nul et normal à la surface en tout point. Par contre, l'image n'en est pas pour autant orientable. En effet, comme le paramétrage n'est pas global, il est possible que plusieurs paramétrages locaux soient nécessaires ; rien ne dit qu'ils indiqueront la même direction.

Si les conditions sont relâchées, le problème persiste.

**Remarque 8.1.5.** Dans un paramétrage régulier<sup>3</sup> le vecteur normal est défini et non-nul en chaque point.

Mais même si l'image est une variété, comme il n'y a pas nécessairement injectivité, il est possible qu'à un même point dans l'image, le paramétrage associe plusieurs vecteurs normaux différents (c'est le cas du ruban de Möbius, exemple 8.2.5).

De plus, l'image n'est pas nécessairement une sous-variété de dimension 2. Le souci est lié à l'injectivité. Par contre, il existe des surfaces paramétrées régulières non-injectives dont l'image est une sous-variété (e.g. le cylindre ou la sphère). Le problème provient plutôt des auto-intersections (comme dans le dessin ci-contre).



Que l'image d'un paramétrage régulier ne soit pas une variété n'est pas non plus automatiquement une conséquence de la non-injectivité : e.g. le cylindre à l'exemple 8.2.4 est une surface qui est aussi l'image d'un paramétrage régulier non-injectif.



Un des aspects pratiques du vecteur normal, c'est qu'il permet de rendre de façon plus concise les informations qui sont utiles pour caractériser un paramétrage régulier.

Par exemple, les propriétés du produit vectoriel permettent entre autres de remarquer qu'un paramétrage  $\Phi$  de classe  $C^1$  est régulier en un point si le vecteur normal ne s'annule pas en ce point ( $\mathbf{N}_\Phi \neq \mathbf{0}$ ).

Avant de présenter une version équivalente, mais plus générale de l'orientabilité, un court préliminaire est requis. En abordant, le travail d'un champ de vecteurs le long d'une courbe paramétrée, il avait été remarqué que cette quantité n'est pas invariante sous reparamétrage. Ceci tient au fait que les reparamétrages  $C^1$  se déclinent en deux catégories : un difféomorphisme  $C^1$  d'un intervalle  $I_1$  dans un autre  $I_2$  possède un signe selon qu'il est croissant ou décroissant.

Les choses en vont de même en dimension supérieure : si  $S$  est un ensemble connexe et  $\mathbf{F} : S \rightarrow \mathbb{R}^n$  est un difféomorphisme de classe  $C^1$  alors  $\text{Det}\mathbf{F}_* : S \rightarrow \mathbb{R}$  est une fonction continue qui ne s'annule pas. Ainsi elle est toujours de même signe.

---

3. Une paramétrage régulier injectif est un paramétrage local... ce qui est important pour cette remarque, est l'absence d'injectivité.



**Proposition 8.1.6**

Soit  $\Phi : D_1 \rightarrow \mathbb{R}^3$  un paramétrage régulier. Soit  $\Psi : D_2 \rightarrow \mathbb{R}^3$  un reparamétrage  $C^1$  obtenu par

composition de  $\Phi$  avec le difféomorphisme  $C^1$ ,  $F: D_2 \rightarrow D_1$ . Alors le signe  $\text{Det } F_*$  détermine si  $N_\Psi$  est positivement ou négativement colinéaire à  $N_\Phi$  :

DÉMONSTRATION: Étant des paramétrages réguliers, les vecteurs normaux sont non-nuls. Il faut seulement déterminer si les vecteurs normaux sont positivement ou négativement colinéaires. Posons  $\mathbf{t} = \mathbf{F}(\mathbf{s})$ , alors par la règle de dérivation en chaîne (théorème 3.4.1)  $\Psi_*(\mathbf{s}) = \Phi_*(\mathbf{t})\mathbf{F}_*(\mathbf{s})$ . Pour  $j \in \overline{2}$ ,  $D_j\Phi(\mathbf{t}) = \Phi_*(\mathbf{t})\mathbf{e}_j$  et  $D_j\Psi(\mathbf{s}) = \Psi_*(\mathbf{s})\mathbf{e}_j = \Phi_*(\mathbf{t})\mathbf{F}_*(\mathbf{s})\mathbf{e}_j$ . Si  $a_{ij}$  sont les coefficients de la matrice  $A = \mathbf{F}_*(\mathbf{s})^\top$ ,

$$\begin{aligned} N_\Psi(\mathbf{s}) &= D_1\Psi(\mathbf{s}) \wedge D_2\Psi(\mathbf{s}) = (a_{11}D_1\Phi(\mathbf{t}) + a_{12}D_2\Phi(\mathbf{t})) \wedge (a_{21}D_1\Phi(\mathbf{t}) + a_{22}D_2\Phi(\mathbf{t})) \\ &= (a_{11}a_{22} - a_{12}a_{21})D_1\Phi(\mathbf{t}) \wedge D_2\Phi(\mathbf{t}) \\ &= (\text{Det } \mathbf{F}_*(\mathbf{s}))N_\Phi(\mathbf{t}) \end{aligned} \quad \blacksquare$$

D'un point de vue intuitif, le résultat ci-dessus n'est peut-être pas surprenant. En effet, le vecteur normal n'est déterminé qu'à un scalaire près, et il n'y a pas beaucoup d'autres possibilités que le déterminant de la matrice de Jacobi pour associer un scalaire à un difféomorphisme  $C^1$  de  $\mathbb{R}^2$ .

## 8.2 L'orientabilité dans $\mathbb{R}^m$

La proposition 8.1.6 permet de passer d'un concept géométrique propre à la dimension 3 (le produit vectoriel et le vecteur normal) en une propriété qui a du sens dans toutes les dimensions (le signe d'un déterminant). En effet, si la surface est recouverte par des paramétrages locaux, chacun d'entre eux donne un vecteur normal. Ces vecteurs normaux sont positivement colinéaires, exactement lorsque le reparamétrage qui relie les deux paramétrages est de déterminant positif.

**Définition 8.2.1.** Si deux paramétrages locaux  $\mathbf{G}_{(i)}: V_i \rightarrow S \cap U_i$  ( $i \in \overline{2}$ ) sont tels que  $S \cap U_1 \cap U_2 \neq \emptyset$ , alors  $\mathbf{T}_{(12)}: V_1 \cap \mathbf{G}_{(1)}^{-1}(S \cap U_2) \rightarrow V_2 \cap \mathbf{G}_{(2)}^{-1}(S \cap U_1)$  définie par  $\mathbf{T}_{(12)} = \mathbf{G}_{(2)}^{-1} \circ \mathbf{G}_{(1)}$  est une application appelée l'application de transition du paramétrage local  $\mathbf{G}_{(1)}$  vers le paramétrage local  $\mathbf{G}_{(2)}$ . ★

### Proposition 8.2.2

Les applications de transition sont des difféomorphismes  $C^1$ . En particulier, si  $S$  est connexe et  $\mathbf{G}_{(i)}: V_i \rightarrow S \cap U_i$  est une famille de paramétrages locaux qui recouvre  $S$  (i.e.  $S \subset \cup_i U_i$ ) alors  $\exists \ell \in \mathbb{Z}_{\geq 1}$  (indépendant de  $i$ ) tel que les  $V_i \subset \mathbb{R}^\ell$ .

Cette dernière proposition peut sembler anodine ou facile. La difficulté provient de ce que la définition de l'application de transition passe par un ensemble qui n'est pas un ouvert (la sous-variété qui est paramétrée). Il faut ainsi passer par un épaississement de cette dernière pour faire une démonstration propre.

Avant de passer au résultat principal, il est bon de noter que la démonstration est identique dans le cadre de "reparamétrages" en dimensions supérieures à 2. Par conséquent, le prochain résultat est une bonne façon de définir l'orientabilité à la fois pour les surfaces dans  $\mathbb{R}^m$  (où  $m > 3$ ) que pour les sous-variétés de dimension plus grande.

**Proposition 8.2.3**

Soit  $S$  une surface et soit donnée une famille d'ouverts  $\{U_i\}_{i \in I}$  (où  $I \subset \mathbb{Z}_{\geq 0}$  est fini ou infini) et de paramétrages locaux  $\Phi_i : V_i \rightarrow S \cap U_i$  (où  $V_i \subset \mathbb{R}^2$  sont des produits d'intervalles).  $S$  est orientable si et seulement si, quitte à inverser certains paramétrages (i.e. remplacer  $\Phi_i$  par  $\Phi_i \circ J$  où  $J(x_1, x_2) = \begin{pmatrix} -x_1 \\ x_2 \end{pmatrix}$ ), les déterminants des matrices de Jacobi des applications  $\mathbf{T}_{(ij)}$  de transition sont positifs,  $\text{Det } \mathbf{T}_{(ij)*} > 0$ .

(Parenthèse : Cette condition signifie qu'en voyant les paramétrages locaux comme des surfaces paramétrées (qui coïncident) et en les mettant bout à bout, le vecteur normal reste toujours du même côté de la surface (c'est-à-dire en un point fixé de la surface tous les vecteurs normaux de tous les paramétrages pointent du même côté) quitte à inverser le vecteur normal de départ pour certains paramétrages.)

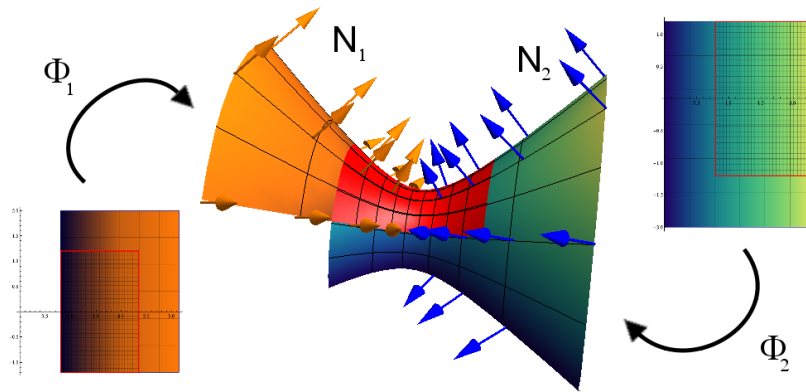
DÉMONSTRATION: Supposons qu'il existe une famille de paramétrages locaux avec la propriété que les applications de transition sont de jacobien positif. Alors ces paramétrages locaux définissent en tout point de leur image un vecteur normal  $D_1\Phi_j \wedge D_2\Phi_j$ . Comme ce sont des paramétrages locaux, ce vecteur n'est jamais nul et varie continûment. Soit

$$\begin{aligned} \mathbf{N}_j : \text{Im } \Phi_j &\rightarrow \mathbb{R}^3 \\ \mathbf{x} &\mapsto \frac{\mathbf{N}_{\Phi_j}(\Phi_j^{-1}(\mathbf{x}))}{\|\mathbf{N}_{\Phi_j}(\Phi_j^{-1}(\mathbf{x}))\|} \end{aligned}$$

ce champ de vecteurs renormalisés pour être unitaire. Soit maintenant  $\Phi_i$  et  $\Phi_j$  deux paramétrages locaux dont les images s'intersectent sur  $W_{ij} = \text{Im } \Phi_i \cap \text{Im } \Phi_j \subset S$  (supposé non-vide). En restreignant le domaine à la préimage de  $W_{ij}$ , ces deux applications sont, via  $\mathbf{T}_{(ij)} = \Phi_j^{-1} \circ \Phi_i$  un reparamétrage  $\mathcal{C}^1$  l'une de l'autre. De plus, comme le déterminant est positif, leur vecteur normal pointera dans la même direction. Cependant, comme il est unitaire, donc  $\mathbf{N}_i = \mathbf{N}_j$  sur  $W_{ij}$ , ou plus rigoureusement, pour  $\mathbf{x} \in W_{ij}$  :

$$\begin{aligned} \mathbf{N}_i(\mathbf{x}) &= \frac{\mathbf{N}_{\Phi_i}(\Phi_i^{-1}(\mathbf{x}))}{\|\mathbf{N}_{\Phi_i}(\Phi_i^{-1}(\mathbf{x}))\|} = \frac{\mathbf{N}_{\Phi_j}(\mathbf{T}_{(ij)} \circ \Phi_i^{-1}(\mathbf{x}))}{\|\mathbf{N}_{\Phi_j}(\mathbf{T}_{(ij)} \circ \Phi_i^{-1}(\mathbf{x}))\|} \frac{\text{Det } \mathbf{T}_{(ij)*}(\mathbf{x})}{|\text{Det } \mathbf{T}_{(ij)*}(\mathbf{x})|} \\ &= \text{sgn}(\text{Det } \mathbf{T}_{(ij)*}(\mathbf{x})) \frac{\mathbf{N}_{\Phi_j}(\Phi_j^{-1}(\mathbf{x}))}{\|\mathbf{N}_{\Phi_j}(\Phi_j^{-1}(\mathbf{x}))\|} = \text{sgn}(\text{Det } \mathbf{T}_{(ij)*}(\mathbf{x})) \mathbf{N}_j(\mathbf{x}) \end{aligned}$$

En particulier, les applications de transition étant de déterminants positifs,  $\mathbf{N}_i = \mathbf{N}_j$  sur  $W_{ij}$ . Ainsi, il est possible de définir un champ de vecteurs  $\mathbf{N} : S \rightarrow \mathbb{R}^3$  continu et qui ne s'annule jamais, en définissant  $\mathbf{N}(\mathbf{x}) = \mathbf{N}_i(\Phi_i^{-1}(\mathbf{x}))$  pour n'importe quel  $\Phi_i$  tel que  $\mathbf{x} \in \text{Im } \Phi_i$ .



Maintenant, commençons par une surface orientable  $S$ , c'ad. soit  $\mathbf{N} : S \rightarrow \mathbb{R}^3$  le champ de vecteurs normaux sur  $S$  (qui sont, sans perte de généralité, supposés unitaires). Il est possible de la couvrir par une famille d'ouverts qui sont les images de paramétrages locaux  $\Phi_i : V_i \rightarrow S$  (et les  $V_i$  peuvent être choisis comme étant connexe par arc, quitte à démultiplier le paramétrage local en le restreignant à différents domaines). Il faut vérifier que ces paramétrages locaux peuvent être corrigés pour que les applications de transition aient des jacobiens positifs.

Comme plus haut, les paramétrages locaux définissent, sur leur image, un champ de vecteurs normaux (unitaires)  $\mathbf{N}_i$ . De ce champ, une nouvelle “fonction” peut être obtenue sur  $\text{Im } \Phi_i$  : la fonction  $E_i : \text{Im } \Phi_i \rightarrow \mathbb{R}$  qui associe à  $\mathbf{x}$  la valeur  $c$  si  $\mathbf{N}_i(\mathbf{x}) = c\mathbf{N}(\mathbf{x})$ .  $E_i$  est une fonction bien définie (car les deux vecteurs sont colinéaires en tout  $\mathbf{x}$ , tous les vecteurs normaux à un plan fixé dans  $\mathbb{R}^3$  le sont), continue et non-nulle (car  $\mathbf{N}$  et  $\mathbf{N}_i$  sont continus et non-nuls), et même identiquement égale à 1 ou  $-1$  (car les deux champs de vecteurs normaux sont unitaires et définis sur un ensemble connexe) ; autrement dit  $E_i = \pm 1$ .

L'idée est alors de changer tous les paramétrages locaux qui ont la mauvaise orientation : si  $\Phi_i$  est un paramétrage local tel que  $E_i = -1$ , alors il est remplacé par  $\Psi_i = \Phi_i \circ J$  où  $J : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  est définie par  $(x, y) \mapsto (-x, y)$ . Comme la matrice de Jacobi de  $J$  (c'ad.  $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ ) est de déterminant négatif, le vecteur normal  $D_1\Psi_i \wedge D_2\Psi_i$  est, une fois renormalisé pour être unitaire, égal à  $\mathbf{N}$  (cf. proposition 8.1.6).

Ainsi, quitte à effectuer ce changement, il est possible de supposer que  $\mathbf{N}_i = \mathbf{N}$ . Soit  $\mathbf{T}_{(ij)}$  une application de transition, elle est définie sur  $W_{ij} \subset V_i$ . Il faut montrer que le jacobien  $\text{Det } \mathbf{T}_{(ij)*}$  est positif sur les composantes connexes de  $W_{ij}$  (le signe est, par continuité et par non-nullité, constant sur les composantes connexes). Comme  $\mathbf{T}_{(ij)}$  est un difféomorphisme  $\mathcal{C}^1$ , il donne un reparamétrage  $\mathcal{C}^1$  de  $\Phi_i$  restreinte à  $W_{ij}$  en  $\Phi_j$  restreinte à  $\mathbf{T}_{(ij)}(W_{ij})$ . Or, si son jacobien est négatif, les vecteurs normaux de  $\Phi_i$  et  $\Phi_j$  doivent être de signes opposés (par la proposition 8.1.6). Donc, le jacobien est positif. ■

Ce qui est important à souligner dans le théorème précédent est que une surface est recouverte par un nombre fini de paramétrages locaux, alors, en un nombre fini d'étapes, il est possible de déterminer si elle est orientable ou non.

**Exemple 8.2.4.** Le cylindre de rayon  $r \in \mathbb{R}_{>0}$ ,  $D = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x^2 + y^2 = r^2 \right\}$ , peut être paramétré

par  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  définie par  $\Phi(s, t) = \begin{pmatrix} r \cos s \\ r \sin s \\ t \end{pmatrix}$ . Le vecteur normal est alors :

$$\begin{pmatrix} -r \sin s \\ r \cos s \\ 0 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -r \sin s & r \cos s & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{pmatrix} r \cos s \\ r \sin s \\ 0 \end{pmatrix}.$$

$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$  est un champ de vecteurs normaux à la surface ; elle est donc orientable. De manière équivalente, deux paramétrages locaux qui permettent de montrer que la surface est orientable sont  $\Phi_1 = \Phi|_{]-\frac{3\pi}{4}, \frac{3\pi}{4}[ \times \mathbb{R}}$  et  $\Phi_2 = \Phi|_{]\frac{\pi}{4}, \frac{7\pi}{4}[ \times \mathbb{R}}$ . ♣

Comme promis, voici un exemple (classique) qui illustre que (dans la proposition 8.2.3) il est important que les paramétrages locaux soient des applications injectives.

**Exemple 8.2.5.** Le ruban de Möbius est l'image du paramétrage décrit par

$$\begin{aligned} \Phi : ]-1, 1[ \times \mathbb{R} &\rightarrow \mathbb{R}^3 \\ (s, t) &\mapsto \begin{pmatrix} (2 + s \cos \frac{t}{2}) \cos t \\ (2 + s \cos \frac{t}{2}) \sin t \\ s \sin \frac{t}{2} \end{pmatrix}. \end{aligned}$$

Voir que c'est une surface est relativement aisé :  $\Phi|_{]-1, 1[ \times ]0, 3\pi/2[}$  et  $\Phi|_{]-1, 1[ \times ]\pi, 5\pi/2[}$  sont toutes les deux des paramétrages locaux et recouvre l'image de  $\Phi$ . Ainsi, chaque point de  $\Phi$  possède un paramétrage local.

En prenant le vecteur normal  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  basé au point  $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ , et en tentant de faire un tour sur cette surface (suivre la courbe  $\Phi(0, t)$ , un cercle de rayon 2) tout en gardant le vecteur toujours du même côté, il apparaît que le vecteur est maintenant de l'autre côté... Cette surface est en effet non-orientable (la démonstration rigoureuse de la non-orientabilité est laissée en exercice, voir ci-après).

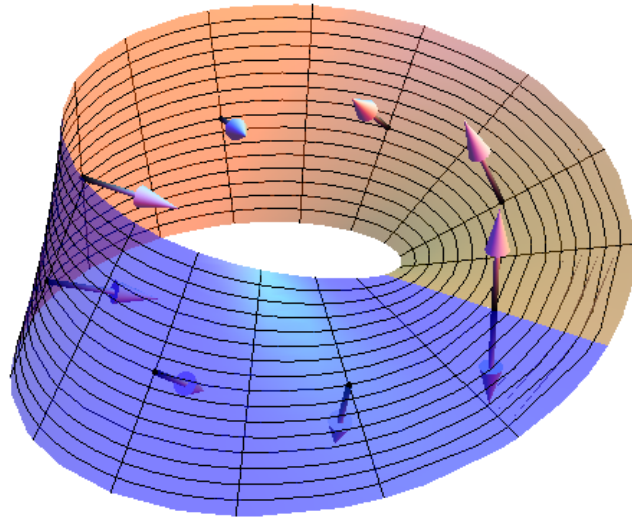
Les calculs se passent comme suit : son vecteur normal est donné par  $D_1\Phi(s, t) \wedge D_2\Phi(s, t) =$

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos \frac{t}{2} \cos t & \cos \frac{t}{2} \sin t & \sin \frac{t}{2} \\ -\frac{s}{2} \sin \frac{t}{2} \cos t - (2 + s \cos \frac{t}{2}) \sin t & -\frac{s}{2} \sin \frac{t}{2} \sin t + (2 + s \cos \frac{t}{2}) \cos t & \frac{s}{2} \cos \frac{t}{2} \end{vmatrix}$$

$$= \begin{pmatrix} \frac{s}{2} \sin t - (2 + s \cos \frac{t}{2}) \sin \frac{t}{2} \cos t \\ -\frac{s}{2} \cos t - (2 + s \cos \frac{t}{2}) \sin \frac{t}{2} \sin t \\ (2 + s \cos \frac{t}{2}) \cos \frac{t}{2} \end{pmatrix}.$$

D'où 
$$\mathbf{N}_\Phi(0, t) = \begin{pmatrix} -2 \sin \frac{t}{2} \cos t \\ -2 \sin \frac{t}{2} \sin t \\ 2 \cos \frac{t}{2} \end{pmatrix}.$$

Ainsi,  $\mathbf{N}_\Phi(0, 0) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$  et (après un tour)  $\mathbf{N}_\Phi(0, 2\pi) = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}$ . En fait, voici une représentation du ruban (le “devant” en orange et le “derrière” en bleu) avec son vecteur normal :



EXERCICE 135: À partir du paramétrage ci-dessus, il est possible de définir deux paramétrages locaux  $\Psi_1 = \Phi|_{]-1,1[ \times ]0, \frac{3\pi}{2}[}$  et  $\Psi_2 = \Phi|_{]-1,1[ \times ]\pi, \frac{5\pi}{2}[}$

a. Sur quel domaine est définie l'application de transition  $T_{(12)}$  ?

b. Déterminer le signe de  $\text{Det } T_{(12)*}$  sur chaque composante connexe de ce domaine.

[Indice : ne pas se lancer dans des calculs avec des fonctions trigonométriques inverses, utiliser que  $\cos(u + \pi) = -\cos u$  et  $\sin(u + \pi) = -\sin u$ .]

EXERCICE 136: Soit  $S \subset \mathbb{R}^3$  est une surface. Supposons qu'il existe deux paramétrages locaux  $\Phi_1$  et  $\Phi_2$  tels que  $\text{Im } \Phi_1 \cup \text{Im } \Phi_2 = S$  et que  $\text{Im } \Phi_1 \cap \text{Im } \Phi_2$  a deux composantes connexes. Si de plus le déterminant de l'application de transition  $T_{(12)}$  est positif sur une composante connexe et négatif sur l'autre, montrer que  $S$  n'est pas orientable.

Les deux exercices précédents montrent que le ruban de Möbius n'est pas orientable.

Pour terminer, voici un exemple plus simple.

**Exemple 8.2.6.** Les cylindres peuvent (à translation et rotation près) être perçus comme les ensembles de niveau (réguliers) associés à la fonction  $G(x, y, z) = x^2 + y^2$  (toutes les valeurs prises par cette fonction, à l'exception notable de 0, sont régulières). Par la proposition 8.1.2, ils sont orientables.

Soit  $r \in \mathbb{R}_{>0}$  fixé et soit  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  le paramétrage défini par  $\Phi(s, t) = \begin{pmatrix} r \cos s \\ r \sin s \\ t \end{pmatrix}$ , alors  $\text{Im } \Phi = G^{-1}(r^2)$ . De plus,  $N_\Phi(s, t) = \begin{pmatrix} r \cos s \\ r \sin s \\ 0 \end{pmatrix}$  n'est jamais nul (régularité du paramétrage) et  $\Phi$  est de classe  $C^\infty$ . Par contre,  $\Phi$  n'est pas injective. Afin de parler en terme de paramétrage local, il suffit ici de prendre des restrictions de  $\Phi$  à un domaine plus petit pour avoir quelque chose d'injectif. Soit

$$\Phi_{(1)} = \Phi|_{]0, 3\pi/2[ \times \mathbb{R}} \quad \text{et} \quad \Phi_{(2)} = \Phi|_{] \pi, 5\pi/2[ \times \mathbb{R}}.$$

Il s'agit de deux paramétrages locaux. Il est évident que  $\Phi_{(1)} \equiv \Phi_{(2)}$  sur  $W = ] \pi, 3\pi/2[ \times \mathbb{R}$ . D'autre part,  $\Phi_{(1)}$  restreinte à  $W_{(1)} = ]0, \pi/2[ \times \mathbb{R}$  a la même image que  $\Phi_{(2)}$  restreinte à  $W_{(2)} = ]2\pi, 5\pi/2[ \times \mathbb{R}$  (puisque  $\Phi$  est  $2\pi$ -périodique sur la première coordonnée). L'application de transition  $T_{(12)}$  est définie sur  $W \cup W_{(1)}$  (et à valeur dans  $W \cup W_{(2)}$ ). Sur  $W$ ,  $T_{(12)}$  est l'identité puisque  $T_{(12)} = \Phi_{(2)}^{-1} \circ \Phi_{(1)}$

et que sur ce domaine les deux applications sont identiques (c'est la même application  $\Phi$  restreint au même domaine). L'application de transition envoie aussi  $W_{(1)}$  vers  $W_{(2)}$ , et (comme l'application  $\Phi$  est  $2\pi$ -périodique) il s'agit que de la translation par  $2\pi$  sur la première coordonnée :  $\begin{pmatrix} s \\ t \end{pmatrix} \mapsto \begin{pmatrix} s+2\pi \\ t \end{pmatrix}$ . D'où,

$$T_{(12)}(s, t) = \begin{cases} \begin{pmatrix} s \\ t \end{pmatrix} & \text{si } s \in ]\pi, 3\pi/2[, \\ \begin{pmatrix} s+2\pi \\ t \end{pmatrix} & \text{si } s \in ]0, \pi/2[. \end{cases}$$

En particulier, la matrice jacobienne de cette application est l'identité et, par conséquent, le jacobien (son déterminant) est positif. (Ceci est une manière compliquée de [re-]démontrer que le cylindre est orientable.) Pour visualiser un peu mieux l'orientation, il est préférable de regarder les vecteurs normaux. ♣

**Exemple 8.2.7.** Une autre (famille de) surface dont l'orientabilité est facilement établie est la famille des sphères (dans  $\mathbb{R}^3$ ). Celles-ci sont (à translation près) les surfaces de niveau de la fonction  $G(x, y, z) = x^2 + y^2 + z^2$ , fonction dont toutes les valeurs (prises) sont régulières, à l'exception de 0. La sphère de rayon  $r \in \mathbb{R}_{>0}$  est  $S = G^{-1}(r^2)$ . Le paramétrage le plus commun est  $\begin{pmatrix} s \\ t \end{pmatrix} \mapsto \begin{pmatrix} r \cos s \sin t \\ r \sin s \sin t \\ r \cos t \end{pmatrix}$ . Cependant, il n'est pas très commode au niveau des possibilités qu'il offre pour donner lieu à des paramétrages locaux (et aux applications de transition associées).

La démarche passera ainsi par les projections stéréographiques (qui sont à proprement parler plutôt des cartes, c'ad. l'application inverse d'un paramétrage local). De ce fait, un rappel de leur construction/définition s'impose. Ce sera fait ici pour la sphère de rayon  $r$  centrée en l'origine ; elle est coupée en deux morceaux égaux par le plan  $z = 0$ . La projection stéréographique (depuis le "nord")  $\mathbf{P}_N : S_r \setminus \left\{ \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix} \right\} \rightarrow \mathbb{R}^2$  est donnée en associant à un point  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in S_r$  différent de  $\begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}$  le point du plan  $z = 0$  qui se situe sur la droite reliant  $\begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}$  à  $p$ . D'où  $\mathbf{P}_N(x, y, z) = \begin{pmatrix} x/(1-z/r) \\ y/(1-z/r) \end{pmatrix}$ . L'autre projection stéréographique, obtenue depuis le pôle sud, c'ad. le point  $\begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix}$ , est donnée par  $\mathbf{P}_S(x, y, z) = \begin{pmatrix} x/(1+z/r) \\ y/(1+z/r) \end{pmatrix}$ .

Pour obtenir des paramétrages locaux, il faut inverser ces applications. Par exemple, pour  $\mathbf{P}_N$  il faut tenter de retrouver, étant donné  $u = \frac{x}{1-z/r}$  et  $v = \frac{y}{1-z/r}$  les trois valeurs  $x, y$ , et  $z$  satisfaisant  $x^2 + y^2 + z^2 = r^2$ . Pour ce faire, remarquons que

$$u^2 + v^2 = r^2 \frac{x^2 + y^2}{(r-z)^2} = r^2 \frac{r^2 - z^2}{(r-z)^2} = r^2 \frac{r+z}{r-z}$$

De là, il apparaît que  $z = r \frac{u^2 + v^2 - r^2}{u^2 + v^2 + r^2}$ . Ensuite, comme  $x = u(1 - \frac{z}{r})$ ,  $x = \frac{2r^2 u}{u^2 + v^2 + r^2}$ , et, de la même façon,  $y = \frac{2r^2 v}{u^2 + v^2 + r^2}$ . D'où :

$$\Phi_N(u, v) = \begin{pmatrix} \frac{2r^2 u}{u^2 + v^2 + r^2} \\ \frac{2r^2 v}{u^2 + v^2 + r^2} \\ \frac{r(u^2 + v^2 - r^2)}{u^2 + v^2 + r^2} \end{pmatrix}$$

est le paramétrage local  $\mathbb{R}^2 \rightarrow S_r \setminus \{(0, 0, r)\}$ . Un calcul similaire donne que

$$\Phi_S(u, v) = \begin{pmatrix} \frac{2r^2 u}{u^2 + v^2 + r^2} \\ \frac{2r^2 v}{u^2 + v^2 + r^2} \\ r \frac{r^2 - u^2 - v^2}{u^2 + v^2 + r^2} \end{pmatrix}$$

est un paramétrage local  $\mathbb{R}^2 \rightarrow S_r \setminus \left\{ \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix} \right\}$ . Puisque  $\Phi_N(0, 0) = \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix}$  et  $\Phi_S(0, 0) = \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}$ , l'application de transition entre ces deux paramétrages locaux est  $\mathbf{T}_{NS} : \mathbb{R}^2 \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \rightarrow \mathbb{R}^2 \setminus \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ . Étant donné que  $\mathbf{T}_{NS} = \Phi_S^{-1} \circ \Phi_N$  et que  $\Phi_S^{-1} = \mathbf{P}_S$ , la formule explicite pour  $\mathbf{T}_{NS}$  est

$$\mathbf{T}_{NS}(u, v) = \begin{pmatrix} \frac{r^2 u}{u^2 + v^2} \\ \frac{r^2 v}{u^2 + v^2} \end{pmatrix}.$$

Il faut ici faire attention : cette application de transition a un jacobien négatif. Cela ne veut pas pour autant dire que la surface n'est pas orientable. De fait, comme le domaine de l'application de transition est fait d'une seule composante connexe, il suffit remplacer  $\Phi_S$  (ou  $\Phi_N$ ) par  $\Phi_S$  composée avec un automorphisme simple du plan qui renverse l'orientation (e.g. l'application linéaire donnée par la matrice  $J = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ ) pour obtenir une nouvelle application dont le jacobien est positif. ♣

EXERCICE 137: Une surface de révolution est paramétrage déterminé par  $f : U \rightarrow \mathbb{R}$  (où  $U \subset \mathbb{R}$ ) une fonction (continue). Alors  $\Phi : U \times \mathbb{R} \rightarrow \mathbb{R}^3$  est défini par  $\Phi(s, t) = \begin{pmatrix} f(s) \cos t \\ f(s) \sin t \\ s \end{pmatrix}$ . [La surface est obtenue en faisant tourner le graphe de  $x = f(z)$  autour de l'axe des  $z$ .]

- Supposons que  $\forall s \in U, f(s) > 0$  et  $I \subset \mathbb{R}$  est un intervalle de longueur strictement inférieure à  $2\pi$ . Soit  $\Psi$  la fonction  $\Phi$  restreinte à  $U \times I$ . Montrer que  $\Psi$  est injective.
- Montrer que si  $\forall s \in U, (f(s)^2)' \neq 0$  alors le paramétrage est régulier.
- Déduire de a. et b. que si  $\forall s \in U, (f(s)^2)' \neq 0$  alors il est possible de trouver deux paramétrages dont l'union des images recouvrent  $\text{Im } \Phi$ .
- Montrer que  $\text{Im } \Phi$  est orientable. [Il est possible d'écrire un champ de vecteurs normaux directement ou d'utiliser les paramétrages.]

### 8.3 Surfaces à bord et orientation induite

Le bord d'un paramétrage  $\Phi : D \rightarrow \mathbb{R}^n$  sera ici définie comme  $\Phi(\partial D)$ . Puisque  $D$  est un ouvert de  $\mathbb{R}^2$  (connexe par arc),  $\Phi$  n'est pas forcément défini sur  $\partial D$ . L'hypothèse implicite de cette section est que  $\Phi$  peut être étendu à  $D'$  un ouvert qui contient  $\overline{D}$ .

Étant donné une courbe paramétrée  $\gamma$  dans  $\mathbb{R}^m$ , il est possible de séparer les reparamétrages en deux classes correspondant au signe du difféomorphisme les reliant à  $\gamma$  (c'est ce qui a été fait en abordant le thème du travail d'un champ de vecteurs le long d'une courbe). Il s'agit essentiellement d'aller dans le même sens que ou dans le sens contraire au premier paramétrage<sup>4</sup>. Ces paramétrages sont alors dits orientés positivement ou négativement.

4. Si  $\gamma$  est injective, c'est même là la seule possibilité pour un paramétrage quelconque.



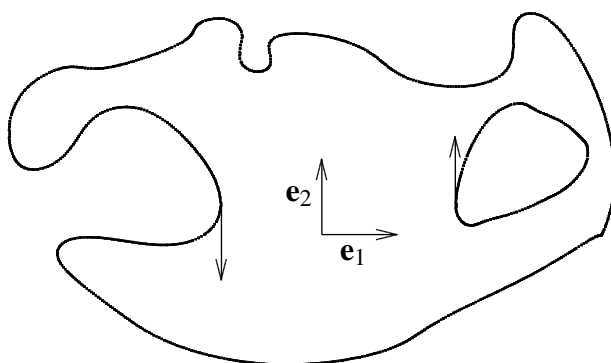
Dans  $\mathbb{R}^2$ , il y a un adage simple pour parler de l'orientation d'un domaine : il faut parcourir le bord de sorte à garder l'intérieur du domaine sur la gauche. Il est d'autant plus raisonnable de s'en tenir à ça puisque seule l'orientation du bord des pavés sera importante. Mais les paragraphes qui suivent en donnent une définition plus rigoureuse (dont l'avantage principal est de mieux comprendre ce qui se passe en dimension supérieure).

Le plus simple est alors de commencer par une orientation de  $\mathbb{R}^2$  (ou  $\mathbb{R}^m$ ). Une orientation est un choix de base orthonormée (où l'ordre est important !). Une telle base peut s'écrire comme une matrice (dite orthogonale). Elles sont soit de déterminant  $+1$  ou  $-1$ . Ainsi, les orientations positives de  $\mathbb{R}^2$  correspondent aux matrices  $R_t := \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$  qui ne sont rien d'autre que la rotation de la base canonique par un angle  $t$ . Les orientations négatives correspondent aux matrices de la forme  $R_t \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .

Dans  $\mathbb{R}^2$  (ou  $\mathbb{R}^m$ ), l'orientation "canonique" ou "standard" est celle qui correspond aux matrices positives, c'ad. à la base canonique  $(\mathbf{e}_{(1)}, \mathbf{e}_{(2)})$  et ses rotations.

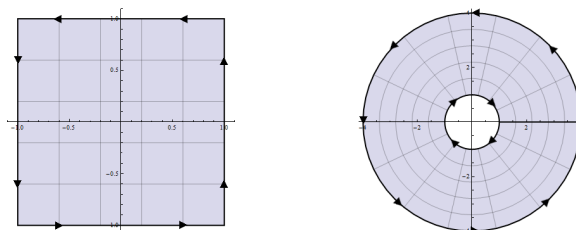
Étant donné  $D$  un ouvert (connexe par arc) de sorte que  $\partial D$  admette un paramétrage par une courbe paramétrée régulière par morceaux, il faut déterminer un sens privilégié pour  $\partial D$ . Comme le bord est une courbe (au sens de variété de dimension 1), il est possible de parler de l'espace tangent à  $x$  dans  $\partial D$  (l'espace engendré par n'importe quel paramétrage régulier en ce point) ; un sous-espace de dimension 1 de  $\mathbb{R}^2$  qui sera noté  $T_x$ .

Soit  $N_x$  l'orthogonal de  $T_x$ , ou autrement dit l'espace normal à  $x$ .  $N_x$  a deux directions, une qui pointe vers l'extérieur et l'autre vers l'intérieur. En  $x \in \partial D$ , il y a une rotation  $R_{\theta_x}$  qui fait en sorte que  $R_{\theta_x} \mathbf{e}_1$  pointe vers l'extérieur. Alors, si  $\gamma : [a, b] \rightarrow \partial D$  paramètre un bout du bord, son orientation est compatible avec l'orientation standard de  $\mathbb{R}^2$  si  $\gamma'(s)$  est positivement colinéaire à  $R_{\theta_x} \mathbf{e}_2$ .



### Exemple 8.3.1.

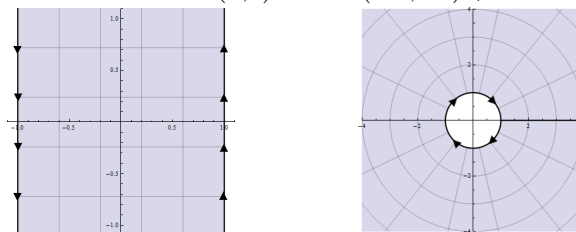
- Le bord du carré  $S = [-1, 1] \times [-1, 1]$  est orienté par  $\gamma(t) = \frac{1}{\max(|\cos t|, |\sin t|)} (\cos t, \sin t)$  (voir ci-dessous).



- Le bord de l'anneau  $S = \{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 4\}$  est formé de deux cercles  $x^2 + y^2 = 1$  et  $x^2 + y^2 = 4$ . L'orientation (i.e. le sens de parcours) de ces deux cercles en tant que bord

de  $S$  est  $t \mapsto (2\cos t, 2\sin t)$  pour le cercle de rayon 2 et  $t \mapsto (\cos t, -\sin t)$  pour celui de rayon 1 (voir ci-dessus).

- Les deux droites  $x = 1$  et  $x = -1$  vues comme le bord de la bande  $S = \{(x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1\}$  obtiennent les orientations  $t \mapsto (1, t)$  et  $t \mapsto (-1, -t)$  (voir ci-dessous).



- Le cercle de rayon 1 vu comme le bord de l'ensemble  $S = \{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2\}$  a l'orientation  $t \mapsto (\cos t, -\sin t)$  (voir ci-dessus).

♣

**Remarque 8.3.2.** Si l'ensemble des traces obtenues en regardant le bord d'ensembles comme ci-dessus est muni d'un signe selon son orientation, les opérations ensemblistes se comportent bien, *c'àd.*

- Si  $D_1 \cap D_2$  est d'intérieur vide, alors  $\partial(D_1 \cup D_2) = \partial D_1 + \partial D_2$  (lorsque qu'un bout de courbe est parcouru une fois dans un sens et une fois dans le sens contraire, ce bout est enlevé).
- Si  $D_2 \subsetneq D_1$  alors  $\partial(D_1 \setminus D_2) = \partial D_1 - \partial D_2$ .

Avec ceci, il est possible de découper la bande en carrés (*cf.* exemple précédent). Puis, chaque carré donne une orientation à son bord. Lorsque l'union des carrés est faite (et une fois retirés les segments parcourus dans un sens et dans l'autre) le bord de la bande obtenue a la même orientation. ♠

Intuitivement, l'orientation du bord de l'image d'un paramétrage régulier est induite par l'orientation de  $\mathbb{R}^2$ . Cependant, ceci ne donne pas nécessairement un bon résultat (regarder ce qui arrive pour le ruban de Möbius). En effet, une courbe est toujours orientable (en particulier, le bord du ruban de Möbius est orientable, même si le ruban ne l'est pas).

En effet, le paramétrage local  $\Phi|_{]-1,1[ \times ]0,2\pi[}$  (où  $\Phi$  est comme à l'exemple 8.2.5) est défini sur un produit d'intervalles ouverts. Si le bord est orienté par une courbe régulière par morceaux  $\gamma$ , alors  $\Phi \circ \gamma$  parcourt bien le bord du ruban de Möbius. Cependant, un segment qui relie  $\Phi(-1, 0)$  avec  $\Phi(1, 0)$  sera parcouru deux fois *dans le même sens*. Autrement dit, l'image du bord de  $D$  par  $\Phi$  ne donne pas le bord du ruban de Möbius.

Finalement, il faut souligner que deux paramétrages locaux d'une même surface qui ont des vecteurs normaux opposés (*i.e.* négativement colinéaires) vont entraîner des orientations opposées des bord.

Revenons à l'exemple 8.2.6 du cylindre.

**Exemple 8.3.3.** À l'exemple 8.2.6, le cylindre était recouvert par deux paramétrages locaux définis sur des pavés "infinis" (un des intervalles est en fait tout  $\mathbb{R}$ ). Il est possible de restreindre lesdites applications encore plus de sorte que l'image soit en fait un cylindre fini (qui aura donc un bord formé de deux cercles) : il suffit de remplacer le  $\Phi$  décrit par  $\Phi|_{\mathbb{R} \times ]a,b[}$  pour obtenir un cylindre de

hauteur  $b - a$ . Autrement dit, de remplacer les deux paramétrages locaux par  $\Phi_{(1)} = \Phi_{]0, 3\pi/2[ \times ]a, b[}$  et  $\Phi_{(2)} = \Phi_{] \pi, 5\pi/2[ \times ]a, b[}$ .

Dans ce cas, les bords du cylindre (les cercles de rayon  $r$  dans les plans  $z = a$  et  $z = b$ ) sont l'image par  $\Phi$  des droites  $\mathbb{R} \times \{a\}$  et  $\mathbb{R} \times \{b\}$ . Pour déterminer, l'orientation de ces cercles, il suffit d'orienter les bords des deux pavés  $]0, 3\pi/2[ \times ]a, b[$  et  $] \pi, 5\pi/2[ \times ]a, b[$  par l'orientation standard de  $\mathbb{R}^2$ , puis de prendre l'image des bouts qui correspondent au bord. Ainsi, il faut regarder la composition de  $\gamma(t) = (t, a)$  (qui lorsque  $t \in [0, 3\pi/2]$  est un paramétrage de  $]0, 3\pi/2[ \times \{a\}$ ) avec  $\Phi_{(1)}$  ainsi que la composition (toujours de)  $\gamma$  (qui lorsque  $t \in ] \pi, 5\pi/2[$  est un paramétrage de  $] \pi, 5\pi/2[ \times \{a\}$ ) avec  $\Phi_{(2)}$  pour avoir le paramétrage du cercle situé sur le plan  $z = a$ . En mettant ces deux bouts de paramétrages du cercle ensemble, le paramétrage  $\hat{\gamma}(t) = \begin{pmatrix} r \cos t \\ r \sin t \\ a \end{pmatrix}$  est l'orientation induite sur ce cercle. Pour l'autre cercle, il est laissé en exercice de vérifier que le paramétrage est  $\hat{\sigma}(t) = \begin{pmatrix} r \cos t \\ -r \sin t \\ b \end{pmatrix}$ .

Finalement, comme auparavant, le calcul du vecteur normal  $\mathbf{N}_\Phi(s, t) = \begin{pmatrix} r \cos s \\ r \sin s \\ 0 \end{pmatrix}$  (en s'écartant de l'axe des  $z$ , ou vers l'"extérieur") rend visuellement bien l'orientation de la surface. ♣

Voici le même exemple traité de manière directe (mais le paramétrage n'est plus fait par un ouvert dans un pavé) et en prenant  $r = 1$ .

**Exemple 8.3.4.** Soit  $a < b \in \mathbb{R}_{>0}$ , soit  $D_{a,b} = \{ \begin{pmatrix} s \\ t \end{pmatrix} \in \mathbb{R}^2 \mid a^2 < s^2 + t^2 < b^2 \}$ , soit  $\Phi : D \rightarrow \mathbb{R}^3$  donnée par

$$\Phi(s, t) = \begin{pmatrix} s/\sqrt{s^2 + t^2} \\ t/\sqrt{s^2 + t^2} \\ \sqrt{s^2 + t^2} \end{pmatrix}.$$

Cette surface paramétrée est un paramétrage (régulier,  $C^1$  et injectif !) possible du cylindre (le produit du cercle de rayon 1 avec l'intervalle  $[a, b]$ ). Comme il n'y a ici qu'une seule paramétrage local (et donc global), la nappe  $S = \text{Im } \Phi$  est automatiquement orientable. En tant que champ de vecteurs normal (puisque l'action se situe dans  $\mathbb{R}^3$ ), cette orientation est donnée par le vecteur normal de  $\Phi$ , soit

$$\mathbf{N}_\Phi(s, t) = D_1 \Phi(s, t) \wedge D_2 \Phi(s, t) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{t^2}{(s^2+t^2)^{3/2}} & \frac{-st}{(s^2+t^2)^{3/2}} & \frac{s}{(s^2+t^2)^{1/2}} \\ \frac{-st}{(s^2+t^2)^{3/2}} & \frac{s^2}{(s^2+t^2)^{3/2}} & \frac{t}{(s^2+t^2)^{1/2}} \end{vmatrix} = \begin{pmatrix} -s/(s^2+t^2) \\ -t/(s^2+t^2) \\ 0 \end{pmatrix}.$$

Ainsi, il s'agit de l'orientation qui pointe vers l'axe des  $z$  (ou vers l'"intérieur", pour autant qu'une telle chose puisse avoir du sens).

Le bord de la nappe  $S = \text{Im } \Phi$  est faite de deux cercles (de rayon 1), un sur le plan  $z = 1$  et un sur le plan  $z = 2$ . Tous deux sont l'image de  $\partial D$  par  $\Phi$ ;  $\partial D$  étant lui-même formé de deux cercles concentriques de rayon  $a$  (le cercle intérieur) et  $b$  (le cercle extérieur). L'orientation canonique de  $\partial D$  est horaire sur le cercle intérieur (e.g. par la courbe paramétrée  $\gamma(t) = \begin{pmatrix} a \cos t \\ -a \sin t \end{pmatrix}$ ) et anti-horaire sur le cercle extérieur (e.g. orienté par la courbe paramétrée  $\sigma(t) = \begin{pmatrix} b \cos t \\ b \sin t \end{pmatrix}$ ). Par conséquent,  $\Phi \circ \gamma(t) = \begin{pmatrix} \cos t \\ -\sin t \\ a \end{pmatrix}$  et  $\Phi \circ \sigma(t) = \begin{pmatrix} \cos t \\ \sin t \\ b \end{pmatrix}$  sont des orientations induites par l'orientation de  $S$  que donne  $\Phi$ .

Ceci est cohérent avec le résultat de l'exemple précédent : une orientation opposée a induit une orientation opposée du bord. ♣

## 8.4 Première forme quadratique : longueur et aire

Afin de commencer en douceur, plutôt que de chercher tout de suite l'aire d'une surface, il vaut mieux regarder un problème qui est essentiellement déjà résolu *c'àd.* la longueur d'une courbe. En effet, si  $\Phi : S \rightarrow \mathbb{R}^m$  est une paramétrage alors à toute courbe paramétrée  $\gamma : [a, b] \rightarrow S$  correspond, par composition, une courbe  $\Phi \circ \gamma$ . Il est alors raisonnable de regarder comment l'application  $\Phi$  transforme les longueurs avant de regarder comment elle transforme l'aire (et peut-être se surprendre à trouver un lien entre les deux).

Ainsi, soit  $\mathbf{x} \in S$ , soit  $\mathbf{v} \in \mathbb{R}^2$  un vecteur, soit  $\gamma : ]-a, a[ \rightarrow S$  telle que  $\gamma(0) = \mathbf{x}$  et  $\gamma'(0) = \mathbf{v}$ . Alors, par le théorème 3.4.1 sur la dérivation en chaîne, si  $\sigma = \Phi \circ \gamma$

$$\sigma'(0) = \Phi_*(\gamma(0))\gamma'(0) = \Phi_*(\mathbf{x})\mathbf{v} = v_1 D_1 \Phi(\gamma(0)) + v_2 D_2 \Phi(\gamma(0)).$$

Ainsi,  $\|\sigma'(0)\|^2 = \|D_1 \Phi(\mathbf{x})\|^2 v_1^2 + 2(D_1 \Phi(\mathbf{x}) \cdot D_2 \Phi(\mathbf{x})) v_1 v_2 + \|D_2 \Phi(\mathbf{x})\|^2 v_2^2$ . La distortion de la distance par  $\Phi$  se mesure par le biais d'une forme quadratique :

$$Q_{\Phi, \mathbf{x}}(\mathbf{v}) = \mathbf{v} \begin{pmatrix} \|D_1 \Phi(\mathbf{x})\|^2 & D_1 \Phi(\mathbf{x}) \cdot D_2 \Phi(\mathbf{x}) \\ D_1 \Phi(\mathbf{x}) \cdot D_2 \Phi(\mathbf{x}) & \|D_2 \Phi(\mathbf{x})\|^2 \end{pmatrix} \mathbf{v}^\top.$$

Cette forme quadratique s'appelle parfois l'élément de longueur ou la première forme fondamentale ; la notation traditionnelle pour cette forme quadratique est  $ds^2$ . Une notation aussi traditionnelle mais qui sera ici utilisée est :

$$\begin{aligned} E_{\Phi}(\mathbf{x}) &= \|D_1 \Phi(\mathbf{x})\|^2, \\ F_{\Phi}(\mathbf{x}) &= D_1 \Phi(\mathbf{x}) \cdot D_2 \Phi(\mathbf{x}), \\ G_{\Phi}(\mathbf{x}) &= \|D_2 \Phi(\mathbf{x})\|^2. \end{aligned} \quad \text{et} \quad Q_{\Phi, \mathbf{x}} = \begin{pmatrix} E_{\Phi}(\mathbf{x}) & F_{\Phi}(\mathbf{x}) \\ F_{\Phi}(\mathbf{x}) & G_{\Phi}(\mathbf{x}) \end{pmatrix}.$$

Comme ce qui a été fait plus haut est général pour toute courbe ayant un vecteur tangent en un point, il s'ensuit que la longueur de  $\Phi \circ \gamma$  est

$$\int_{\Phi \circ \gamma} 1 = \int_a^b \sqrt{Q_{\Phi, \gamma(t)}(\gamma'(t))} dt.$$

Il n'est pas nécessaire de mettre une valeur absolue sous la racine puisque  $Q_{\Phi, \mathbf{x}}$  est toujours semi-définie positive. Effectivement, par définition  $Q_{\Phi, \mathbf{x}}(\mathbf{v})$  est la longueur d'un vecteur.

**Exemple 8.4.1.** Soit  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  le paramétrage défini par  $\Phi(s, t) = \begin{pmatrix} \cos s \sin t \\ \sin s \sin t \\ \cos t \end{pmatrix}$  (et décrivant une sphère de rayon 1). D'où

$$D_1 \Phi(s, t) = \begin{pmatrix} -\sin s \sin t \\ \cos s \sin t \\ 0 \end{pmatrix} \text{ et } D_2 \Phi(s, t) = \begin{pmatrix} \cos s \cos t \\ \sin s \cos t \\ -\sin t \end{pmatrix}.$$

De là,

$$E_{\Phi}(s, t) = \sin^2 t, \quad F_{\Phi}(s, t) = 0, \quad \text{et } G_{\Phi}(s, t) = 1.$$

Ce qui permet de trouver la longueur de courbes sur la sphère.



Pour mesurer l'aire, l'idée est essentiellement la même que pour la longueur, *c'àd.* de séparer l'ensemble  $S$  en petits carrés et de regarder l'aire de l'image de chaque petit carré. En effet, au point  $\mathbf{x} \in S$  le rectangle  $[x_1, x_1 + \varepsilon] \times [x_2, x_2 + \delta]$  (d'aire  $\varepsilon\delta$ ) sera envoyé vers un parallélogramme dont trois sommets seront (approximativement)  $\Phi(\mathbf{x})$ ,  $\Phi(\mathbf{x}) + \varepsilon D_1 \Phi(\mathbf{x})$  et  $\Phi(\mathbf{x}) + \delta D_2 \Phi(\mathbf{x})$ . Or l'aire de ce parallélogramme est la norme du vecteur  $\varepsilon D_1 \Phi(\mathbf{x}) \wedge \delta D_2 \Phi(\mathbf{x}) = \varepsilon\delta \mathbf{N}_\Phi(\mathbf{x})$ . Autrement dit, l'aire est (infinitésimalement en  $\mathbf{x}$ ) multipliée par  $\|\mathbf{N}_\Phi(\mathbf{x})\|$ .

Il est alors naturel de penser à l'aire comme  $\int_S \|\mathbf{N}_\Phi\|$ .

Cette définition ne semble pas faire intervenir les mêmes ingrédients que la mesure de la longueur, ce qui est pour le moins surprenant. Il suffit en fait de creuser un peu (*cf.* proposition 1.4.1) pour remarquer que

$$\begin{aligned} \|\mathbf{N}_\Phi(\mathbf{x})\|^2 &= \|D_1 \Phi(\mathbf{x}) \wedge D_2 \Phi(\mathbf{x})\|^2 = \|D_1 \Phi(\mathbf{x})\|^2 \|D_2 \Phi(\mathbf{x})\|^2 - (D_1 \Phi(\mathbf{x}) \cdot D_2 \Phi(\mathbf{x}))^2 \\ &= E_\Phi(\mathbf{x}) G_\Phi(\mathbf{x}) - F_\Phi(\mathbf{x})^2 \\ &= \text{Det } Q_{\Phi, \mathbf{x}}. \end{aligned}$$

Ce passage n'est pas qu'une pure consolation mathématique. En fait, l'avantage très supérieur de  $Q_\Phi$  sur  $\mathbf{N}_\Phi$  est que la première est définie pour un paramétrage dans  $\mathbb{R}^m$  pour n'importe quel  $m \geq 2$  tandis que le second n'est défini que dans  $\mathbb{R}^3$ .

La définition formelle sera faite pour les intégrales de fonction, en se rappelant que l'aire est obtenue en prenant la fonction constante de valeur 1.

**Définition 8.4.2.** Soit  $H : \mathbb{R}^m \rightarrow \mathbb{R}$  une fonction. L'intégrale de  $H$  sur la surface paramétrée  $\Phi : S \rightarrow \mathbb{R}^m$  est définie par

$$\int_\Phi H = \int_S H(\Phi(\mathbf{x})) \sqrt{\text{Det } Q_{\Phi, \mathbf{x}}} d\text{vol}(\mathbf{x}). \quad \diamond$$

★

En particulier, si  $H \equiv 1$ , il s'agit alors de l'aire de  $\Phi$ . Notons que selon le comportement de  $H$  ou de  $\mathbf{N}_\Phi$  la définition ci-dessus peut donner lieu à une intégrale impropre. Lorsque  $\Phi$  est bijective sur  $\mathring{S}$ , l'aire de  $\Phi : S \rightarrow \mathbb{R}^m$  est donnée par

$$\int_\Phi 1 = \int_S \|\mathbf{N}_\Phi(\mathbf{x})\| d\text{vol}(\mathbf{x}) = \int_S \sqrt{E_\Phi(\mathbf{x}) G_\Phi(\mathbf{x}) - F_\Phi(\mathbf{x})^2} d\text{vol}(\mathbf{x}) = \int_S \sqrt{\text{Det } Q_{\Phi, \mathbf{x}}} d\text{vol}(\mathbf{x}).$$

[La quantité sous la racine est toujours positive ou nulle, sinon il y a erreur de calcul.] Avant de faire quelques exemples, montrons que la définition 8.4.2 est bien invariante par reparamétrage  $C^1$ . Pour ce faire, il faut d'abord remonter un petit résultat sur la transformation que subit la première forme fondamentale lorsqu'il y a reparamétrage.

#### Lemme 8.4.3

Soit  $\Phi : D_1 \rightarrow \mathbb{R}^m$  un paramétrage. Soit  $\Psi : D_2 \rightarrow \mathbb{R}^m$  un reparamétrage  $C^1$  obtenu par composition de  $\Psi$  avec le difféomorphisme  $C^1$ ,  $F : D_2 \rightarrow D_1$ . Alors

$$\text{Det } Q_\Psi(s) = \text{Det } F_*(s)^2 \text{Det } Q_\Phi(F(s)).$$

DÉMONSTRATION: Comme pour la proposition 8.1.6, posons  $\mathbf{t} = \mathbf{F}(\mathbf{s})$ , alors par la règle de dérivation en chaîne (théorème 3.4.1)  $\Psi_*(\mathbf{s}) = \Phi_*(\mathbf{t})\mathbf{F}_*(\mathbf{s})$ . Pour  $j \in \overline{2}$ ,  $D_j\Phi(\mathbf{t}) = \mathbf{e}_j\Phi_*(\mathbf{t})^\top$  et  $D_j\Psi(\mathbf{s}) = \mathbf{e}_j\Psi_*(\mathbf{s})^\top = \mathbf{e}_j\mathbf{F}_*(\mathbf{s})^\top\Phi_*(\mathbf{t})^\top$ . Si  $a_{ij}$  sont les coefficients de la matrice  $A = \mathbf{F}_*(\mathbf{s})^\top$ ,

$$\begin{aligned} E_\Psi(\mathbf{s}) &= \|D_1\Psi(\mathbf{s})\|^2 = \|a_{11}D_1\Phi(\mathbf{t}) + a_{12}D_2\Phi(\mathbf{t})\|^2 = a_{11}^2E_\Phi(\mathbf{t}) + a_{11}a_{12}F_\Phi(\mathbf{t}) + a_{12}^2G_\Phi(\mathbf{t}) \\ F_\Psi(\mathbf{s}) &= D_1\Psi(\mathbf{s}) \cdot D_2\Psi(\mathbf{s}) = a_{11}a_{21}E_\Phi(\mathbf{t}) + (a_{11}a_{22} + a_{12}a_{21})F_\Phi(\mathbf{t}) + a_{12}a_{22}G_\Phi(\mathbf{t}) \\ G_\Psi(\mathbf{s}) &= \|D_2\Psi(\mathbf{s})\|^2 = \|a_{21}D_1\Phi(\mathbf{t}) + a_{22}D_2\Phi(\mathbf{t})\|^2 = a_{21}^2E_\Phi(\mathbf{t}) + a_{21}a_{22}F_\Phi(\mathbf{t}) + a_{22}^2G_\Phi(\mathbf{t}) \end{aligned}$$

Ainsi,  $E_\Psi(\mathbf{s})G_\Psi(\mathbf{s}) - F_\Psi(\mathbf{s})^2 = (a_{11}a_{22} - a_{12}a_{21})^2(E_\Phi(\mathbf{t})G_\Phi(\mathbf{t}) - F_\Phi(\mathbf{t})^2)$ , comme attendu. ■

#### Proposition 8.4.4

Soit  $D_1$  un ouvert borné, soit  $\Phi : D_1 \rightarrow \mathbb{R}^m$  un paramétrage régulier et  $\Psi : D_2 \rightarrow \mathbb{R}^m$  un paramétrage équivalent à  $\Phi$ , alors si  $F$  est une fonction continue sur  $\text{Im } \Phi$ ,

$$\int_\Phi F = \int_\Psi F.$$

DÉMONSTRATION: La démonstration est essentiellement une reprise de celle de la propositions 4.3.9 et 8.4.3. En effet, soit  $\mathbf{H} : D_2 \rightarrow D_1$  le difféomorphisme  $\mathcal{C}^1$  tel que  $\Psi = \Phi \circ \mathbf{H}$ . Alors comme à la proposition 8.4.3,  $\sqrt{\text{Det } Q_\Psi(\mathbf{x})} = |\text{Det } \mathbf{H}_*(\mathbf{x})| \sqrt{\text{Det } Q_\Phi(\mathbf{H}(\mathbf{x}))}$ .

$$\begin{aligned} \int_\Phi F &= \int_{D_1} F(\Phi(\mathbf{y})) \sqrt{\text{Det } Q_\Phi(\mathbf{y})} d\text{vol}(\mathbf{y}) = \int_{\mathbf{H}(D_2)} F(\Phi(\mathbf{y})) \sqrt{\text{Det } Q_\Phi(\mathbf{y})} d\text{vol}(\mathbf{y}) \\ &= \int_{D_2} |\text{Det } \mathbf{H}_*(\mathbf{x})| F(\Phi \circ \mathbf{H}(\mathbf{y})) \sqrt{\text{Det } Q_\Phi(\mathbf{H}(\mathbf{y}))} d\text{vol}(\mathbf{y}) = \int_{D_2} F(\Psi(\mathbf{y})) \sqrt{\text{Det } Q_\Psi(\mathbf{y})} d\text{vol}(\mathbf{y}) \\ &= \int_\Psi F, \end{aligned}$$

par application de la formule du changement de variable 7.4.2. ■

Pour parler d'intégrale sur une surface qui ne possède pas de paramétrage global, il faut procéder par des découpages. Ceci complique beaucoup les notations et les énoncés sans apporter beaucoup, et ne sera pas fait ici.

En bref, étant donné une famille de paramétrage  $\Phi_i : U_i \rightarrow S$ , il faut regarder des sous-ensembles  $D_i$  de chaque  $U_i$  de sorte que : 1- chaque point de la surface apparienne au plus à un des  $\Phi_i(D_i)$ , 2- la préimage des points qui ne sont pas dans un des  $\Phi_i(D_i)$  est un ensemble négligeable dans chaque  $D_j$ .

En particulier, si il y a un paramétrage local dont l'image ne manque que quelques points ou courbes (lisses), il suffit de faire l'intégrale pour ce paramétrage.

**Exemple 8.4.5.** Le tore (de rayons  $r$  et  $R$ ) est une surface qui peut être vue comme l'image du paramétrage  $\Phi : ]0, 2\pi[ \times ]0, 2\pi[ \rightarrow \mathbb{R}^3$  défini par  $\Phi(s, t) = \begin{pmatrix} (R+r\cos s)\cos t \\ (R+r\cos s)\sin t \\ r\sin s \end{pmatrix}$ . Les points qui ne sont pas dans l'image sont dans l'union de deux cercles, un ensemble négligeable (pour une intégrale en dimension 2). Le vecteur normal pour ce paramétrage est

$$\begin{aligned} \mathbf{N}_\Phi(s, t) &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -r\sin s \cos t & -r\sin s \sin t & r\cos s \\ -(R+r\cos s)\sin t & (R+r\cos s)\cos t & 0 \end{vmatrix} \\ &= -r(R+r\cos s) \begin{pmatrix} \cos t \cos s \\ \sin t \cos s \\ \sin s \end{pmatrix} \end{aligned}$$

Ainsi  $\|\mathbf{N}_\Phi(s, t)\| = \sqrt{\text{Det } Q_\Phi(s, t)} = r|R + r \cos s|$ . En supposant que  $0 \leq r \leq R$ , il apparaît que

$$\int_\Phi 1 = r \int_0^{2\pi} \int_0^{2\pi} (R + r \cos s) \, d\mathbf{s} = 4\pi Rr.$$

Autrement dit, l'aire de ce tore est  $4\pi Rr$ . ♣

La physique associe aussi une quantité importante à un champ de vecteurs le long d'une surface. L'idée est d'alors voir  $\mathbb{R}^3$  rempli d'un fluide et le champ de vecteurs comme (à un instant donné) la quantité/vitesse de/du fluide qui se déplace dans la direction donnée. Le flux à travers une surface représente alors la quantité de fluide qui traverse la surface. Il faut évidemment fixer une orientation pour pouvoir parler de ce concept, c'ad. comme le travail, la définition dépendra à un signe près de l'orientation.

**Définition 8.4.6.** Soit  $\Phi : D \rightarrow \mathbb{R}^3$  un paramétrage local, soit  $U$  un ouvert contenant la surface  $\text{Im } \Phi$ , et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^3$  un champ de vecteurs continu et borné. Alors le flux de  $\mathbf{F}$  à travers  $\text{Im } \Phi$  est défini par

$$\int_\Phi \mathbf{F} \cdot d\mathbf{N} = \int_D \mathbf{F}(\Phi(s, t)) \cdot \mathbf{N}_\Phi(s, t) \, d\text{vol}(s, t).$$

★

EXERCICE 138: Montrer que si  $\Phi$  est un reparamétrage  $C^1$  de  $\Psi$ , alors le flux de  $\mathbf{F}$  par  $\Phi$  est à un signe près le flux de  $\mathbf{F}$  par  $\Psi$  et que ce signe est le signe du jacobien du difféomorphisme de reparamétrage.

EXERCICE 139: Montrer que le flux à travers un ruban de Möbius n'est pas invariant par paramétrage, c'ad. trouver un champ de vecteurs et deux paramétrages (dont le complémentaire de l'image est négligeable) de sorte que la valeur absolue du flux est différente.

**Exemple 8.4.7.** Le champ électrique est donné par un champ de vecteurs. Une charge électrique ponctuelle produit un champ  $\mathbf{F} : \mathbb{R}^3 \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}^3$  donné par  $\mathbf{F}(\mathbf{x}) = k \frac{\mathbf{x}}{\|\mathbf{x}\|^3}$  où  $k \in \mathbb{R}_{>0}$ . Cherchons le flux de ce champ à travers une sphère de rayon  $r$  orientée vers l'intérieur (l'origine). Pour ce faire, le paramétrage choisi est  $\Phi : [-\pi/2, \pi/2] \times [0, 2\pi] \rightarrow \mathbb{R}^3$  défini par  $\Phi(s, t) = \begin{pmatrix} r \cos s \cos t \\ r \cos s \sin t \\ r \sin s \end{pmatrix}$ . De nouveau, il s'agit d'un découpage en 1 morceau. Tout d'abord le vecteur normal se calcule par

$$\mathbf{N}_\Phi(s, t) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -r \sin s \cos t & -r \sin s \sin t & r \cos s \\ -r \cos s \sin t & r \cos s \cos t & 0 \end{vmatrix} = -r^2 \cos s \begin{pmatrix} \cos s \cos t \\ \cos s \sin t \\ \sin s \end{pmatrix}$$

Comme  $\mathbf{F}(\Phi(s, t)) = \frac{k}{r^2} \begin{pmatrix} \cos s \cos t \\ \cos s \sin t \\ \sin s \end{pmatrix}$ , le flux est

$$\int_\Phi \mathbf{F} \cdot d\mathbf{N} = \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} -k \cos s \, d\mathbf{s} = -4\pi k.$$

En particulier, le flux est constant quel que soit le rayon choisi. ♣

EXERCICE 140: Soit  $U$  la boule (ouverte) de rayon 1 centrée en  $\mathbf{0} \in \mathbb{R}^2$ , et soit  $\Phi : U \rightarrow \mathbb{R}^3$  le paramétrage défini par  $\Phi(s, t) = \begin{pmatrix} s \\ t \\ \sqrt{1-s^2-t^2} \end{pmatrix}$ . Calculer sa première forme fondamentale. En déduire l'aire de l'image de  $\Phi$ . Expliquer intuitivement le résultat obtenu.

EXERCICE 141: Soit  $\gamma : [0, L] \rightarrow \mathbb{R}^2$  une courbe paramétrée par la longueur d'arc,  $\gamma(s) = \begin{pmatrix} \gamma_1(s) \\ 0 \\ \gamma_2(s) \end{pmatrix}$ .

Il est possible d'obtenir une surface de révolution (plus générale que celles décrites précédemment) en faisant la rotation de cette courbe autour du 3<sup>ème</sup> axe. Soit  $\Phi : ]0, L[ \times \mathbb{R} \rightarrow \mathbb{R}^3$  défini par

$$\Phi(s, t) = \begin{pmatrix} \gamma_1(s) \cos t \\ \gamma_1(s) \sin t \\ \gamma_2(s) \end{pmatrix}.$$

- Si  $\gamma$  est injective, l'image de  $\Phi$  est-elle une surface (au sens de variété de dimension 2) ? Si oui, est-elle orientable ?
- Donner la première forme fondamentale de  $\Phi$ .
- Montrer que si  $p(s)$  est la distance de  $\gamma(s)$  à l'axe de rotation, alors, l'aire de l'image par  $\Phi$  de  $]0, L[ \times ]0, 2\pi[$  est

$$2\pi \int_0^L p(s) ds$$

- Utiliser ceci pour calculer l'aire d'un tore.

EXERCICE 142: Soit  $U \subset \mathbb{R}^2$  un ouvert connexe par arc et  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . Soit  $S$  l'ensemble (dans  $\mathbb{R}^3$ ) formé par le graphe de  $D$ . Le paramétrage le plus simple de cette surface est  $\Phi : U \rightarrow S$  défini par  $\Phi(x, y) = \begin{pmatrix} x \\ y \\ F(x, y) \end{pmatrix}$ .

- Montrer que l'image de  $\Phi$  est une surface (au sens de variété de dimension 2).
- La surface est-elle orientable ?
- Montrer que sa première forme fondamentale est

$$Q_{\Phi, (x, y)} = \begin{pmatrix} 1 + D_1 F(x, y)^2 & D_1 F(x, y) D_2 F(x, y) \\ D_1 F(x, y) D_2 F(x, y) & 1 + D_2 F(x, y)^2 \end{pmatrix}.$$

EXERCICE 143: Soit  $U \subset \mathbb{R}^3$  un ouvert, soit  $F : U \rightarrow \mathbb{R}$  une fonction et  $S_0 = F^{-1}(0)$ . Supposons que 0 est une valeur régulière de  $F$ . Soit  $\pi_{12} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  définie par  $\pi_{12}(x, y, z) = \begin{pmatrix} x \\ y \end{pmatrix}$ , et supposons de plus que  $\pi_{12}$  est injective sur  $S_0$ .

- Pourquoi existe-t-il une application de classe  $C^0$  définie sur un ouvert  $V \subset \mathbb{R}^2$ , telle que  $\Phi : V \rightarrow S_0$  est bijective ?
- Montrer que si  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in S_0$  est tel que  $D_3 F(x, y, z) \neq 0$  il existe un voisinage  $V' \subset V$  de  $\begin{pmatrix} x \\ y \end{pmatrix}$  sur lequel  $\Phi$  est un paramétrage local.
- Montrer que la première forme fondamentale pour ce paramétrage local  $\Phi$  s'exprime par

$$Q_{\Phi, (x, y)} = \begin{pmatrix} 1 + \frac{D_1 F(x, y, z)^2}{D_3 F(x, y, z)^2} & \frac{D_1 F(x, y, z) D_2 F(x, y, z)}{D_3 F(x, y, z)^2} \\ \frac{D_1 F(x, y, z) D_2 F(x, y, z)}{D_3 F(x, y, z)^2} & 1 + \frac{D_2 F(x, y, z)^2}{D_3 F(x, y, z)^2} \end{pmatrix}$$



où  $z$  dépend implicitement de  $\begin{pmatrix} x \\ y \end{pmatrix}$ .

EXERCICE 144: Soit  $\Phi : [0, 1]^2 \rightarrow \mathbb{R}^3$  la surface déterminée par  $\Phi(r, s) = \begin{pmatrix} r^2 \cos s \\ r^2 \sin s \\ r^4/2 \end{pmatrix}$ .

- Calculer le vecteur normal. Pour quelles valeurs de  $\begin{pmatrix} r \\ s \end{pmatrix}$  est-il nul ?
- En déduire que  $\Phi$  est régulière sur l'intérieur de son domaine (c'ad.  $]0, 1[^2$ ). Quelle est l'aire de  $\Phi$  ?
- Soit  $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  le champ de vecteurs défini par  $F(x, y, z) = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$ . Quel est le flux  $\int_{\Phi} F \cdot dN$  de  $F$  à travers  $\Phi$  ?

EXERCICE 145: La caténoïde peut être paramétrée par l'application  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  définie par  $\Phi(s, t) = \begin{pmatrix} \cosh s \cos t \\ \cosh s \sin t \\ s \end{pmatrix}$ .

- Ce paramétrage est-il injectif ? est-il  $C^1$  ?
- Calculer le vecteur normal  $N_{\Phi}$ . En déduire les points où le paramétrage est régulier. L'image de  $\Phi$  est-elle une surface ?
- Quelle est l'aire de l'image par  $\Phi$  du domaine  $] -1, 1[ \times ]0, 2\pi[$  ?
- Cette surface est-elle orientable ?

EXERCICE 146: Une des possibilités pour décrire un cône est le paramétrage suivant :  $\Phi(u, v) = \begin{pmatrix} u \cos v \\ u \sin v \\ u \end{pmatrix}$  défini pour  $u \in \mathbb{R}_{>0}$  et  $v \in ]0, 2\pi[$  et à valeur dans  $\mathbb{R}^3$ .

- Montrer que la matrice de la première forme fondamentale de  $\Phi$  en  $\begin{pmatrix} u \\ v \end{pmatrix}$  est  $Q_{\Phi, (u, v)} = \begin{pmatrix} 2 & 0 \\ 0 & u^2 \end{pmatrix}$ , c'ad. montrer que  $E_{\Phi}(u, v) = 2$ ,  $G_{\Phi}(u, v) = u^2$  et que  $F_{\Phi}(u, v) = 0$ .
- Soit donnée la courbe paramétrée  $\gamma : ]0, \pi/2[ \rightarrow \mathbb{R}_{>0} \times ]0, 2\pi[$  définie par  $\gamma(t) = \begin{pmatrix} \cos t \\ \sqrt{2}t \end{pmatrix}$ . Trouver la longueur de la courbe  $\sigma = \Phi \circ \gamma$ . [Indice : se servir de la première forme fondamentale.]
- Soit  $\Psi = \Phi|_{]1, 2[ \times ]0, 2\pi[}$  le paramétrage ci-haut mais restreint à  $]1, 2[ \times ]0, 2\pi[$ . Quelle est l'aire de l'image de  $\Psi$  ?

## 8.5 Deuxième forme fondamentale : courbure

Tout comme pour les courbes, il est possible (et même souhaitable) de parler de la courbure d'une surface. Par exemple, les formes que dessine un film de savon tendu entre des structures métalliques sont caractérisées par le fait que la surface est minimale. Cette propriété se traduit en terme de courbure par le fait que la courbure moyenne est nulle.

D'autres phénomènes importants sont liés à la courbure de Gauss, e.g. étant donné deux courbes (parcourues à vitesse constante) qui réalisent un plus court chemin sur la surface et se croisent en un point avec un angle  $\theta$ , à quelle vitesse ces deux trajectoires divergent ? Dans le plan, par exemple, ces courbes sont des lignes droites et l'écart est linéaire. Sur une sphère, il est moins facile de voir de quoi ont l'air ces courbes mais, vu que la sphère est compacte, l'écart ne peut pas toujours augmenter.

Un autre phénomène qu'il est assez facile de se convaincre est différent sur une sphère que sur un plan est la somme des angles d'un triangle. Sur un plan, cette somme est  $\pi$ . Sur sphère, en prenant

par exemple  $(0^\circ N, 10^\circ E)$  (pas très loin de Libreville au Gabon),  $90^\circ N$  (le pôle nord) et  $(0^\circ N, 80^\circ O)$  (pas très loin de Bahía de Caráquez en Équateur), un triangle dont tous les angles sont à  $90^\circ$ , ce qui donne une somme de  $3\pi/2$ ... [Il faut savoir que les morceaux d'équateur et de méridiens sont des chemins les plus courts sur la sphère.]

La manière la plus simple de parler de courbure sur une sphère est de commencer par la courbure d'une courbe. Une première tentative de définition serait :

“Soit  $\gamma : [-1, 1] \rightarrow S$  une courbe qui prend valeur sur une surface  $S$  et soit  $\mathbf{x} = \gamma(0) \in S$ . Supposons que  $\gamma$  est paramétrée par la longueur d'arc et que la courbure de  $\gamma$  (représentée par  $\gamma''$ ) soit uniquement dû au fait que  $\gamma$  doit rester dans la surface. Alors, la courbure de  $S$  dans la direction  $\gamma'(0)$  en  $\mathbf{x}$  est la courbure de  $\gamma$ .”

Deux remarques importantes :

- il est possible de parler de courbure sans parler de paramétrage par la longueur d'arc, voir exercice 81 au chapitre 4.
- si une partie de la courbure de  $\gamma$  (en 0, avec  $\gamma(0) = \mathbf{x}$ ) ne provient pas du fait que  $\gamma$  reste dans  $S$ , alors cette partie de la courbure est la projection du vecteur de courbure dans l'espace tangent  $T_{\mathbf{x}}S$ .

Ainsi, une seconde tentative de définition est :

“Soit  $\gamma : [-1, 1] \rightarrow S$  une courbe qui prend valeur sur une surface  $S$  et soit  $\mathbf{x} = \gamma(0) \in S$ . La courbure de  $S$  dans la direction  $\gamma'(0)$  en  $\mathbf{x}$  est la partie de la courbure de  $\gamma$  qui est orthogonale à  $T_{\mathbf{x}}S$ .”

Il n'est cependant pas clair que ceci est bien défini (*i.e.* que ça ne dépende que de  $\gamma'(0)$  et non de  $\gamma''(0)$  ou d'autres paramètres)... Un petit calcul est alors de propos. Soit  $\Phi : D \rightarrow S$  un paramétrage local (régulier suffirait) et  $\sigma : [-1, 1] \rightarrow D$  une courbe. Pour alléger les notations, soient

$$\sigma(0) = \mathbf{s}, \quad \mathbf{v} = \sigma'(0), \quad \Phi(\mathbf{s}) = \mathbf{x} \quad \text{et} \quad V \text{ l'orthogonal de } T_{\mathbf{s}}^{\text{par}} \Phi = T_{\mathbf{x}}^{\text{vit}} S$$

Alors le vecteur de courbure de  $\gamma = \Phi \circ \sigma$  est (par l'exercice 81) :

$$\mathbf{k}(t) = \frac{\|\gamma'(t)\|^2 \gamma''(t) - (\gamma'(t) \cdot \gamma''(t)) \gamma'(t)}{\|\gamma'(t)\|^4}$$

Mais

$$\begin{aligned} \gamma'(0) &= D_1 \Phi(\mathbf{s}) v_1 + D_2 \Phi(\mathbf{s}) v_2 \\ \text{et } \gamma''(0) &= D_1 \Phi(\mathbf{s}) \sigma''(0) + D_2 \Phi(\mathbf{s}) \sigma''(0) + \sum_{i=1}^2 \sum_{j=1}^2 D_j D_i \Phi(\mathbf{s}) v_i v_j \end{aligned}$$

Ainsi,  $\mathbf{k}(t)$  est une combinaison linéaire de  $\gamma'(0)$  et  $\gamma''(0)$ . Cependant, seule la partie de  $\mathbf{k}$  qui est orthogonale à l'espace tangent doit être évaluée. Or l'espace tangent en  $\mathbf{x}$  est engendré par  $D_1 \Phi(\mathbf{s})$  et  $D_2 \Phi(\mathbf{s})$ . Puisque la quantité d'intérêt est la projection de  $\mathbf{k}(0)$  sur  $V$ , tous les termes qui contiennent  $D_1 \Phi(\mathbf{s})$  ou  $D_2 \Phi(\mathbf{s})$  peuvent être retirés (*e.g.* les  $\gamma'(0)$ ). Il ne reste alors qu'à trouver la projection sur  $V$  de

$$\tilde{\mathbf{k}}(0) := \frac{\sum_{i=1}^2 \sum_{j=1}^2 D_j D_i \Phi(\mathbf{s}) v_i v_j}{\mathbf{v}^T Q_{\Phi, \mathbf{s}} \mathbf{v}};$$

puisque  $\|\gamma'(t)\|^2 = \mathbf{v}^T Q_{\Phi, \mathbf{s}} \mathbf{v}$  où  $Q_{\Phi, \mathbf{s}}$  est la première forme fondamentale. Ce qui est déjà étonnant (et réconfortant) est que le vecteur ci-dessus ne dépend pas de  $\sigma$  mais uniquement de  $\mathbf{v} = \sigma'(0)$ . Ainsi, pour n'importe quelle courbe qui aura le même vecteur tangent, le résultat sera identique. D'où

**Théorème 8.5.1 (Meusnier)**

Toutes les courbes qui passe par un point d'une surface avec le même vecteur tangent ont la même courbure normale à la surface.

Pour simplifier la suite du calcul, supposons que  $m = 3$ . Alors,  $V$  est l'espace engendré par  $\mathbf{N}_\Phi$  et il suffit de prendre le produit scalaire avec  $\mathbf{N}_\Phi / \|\mathbf{N}_\Phi\|$  pour trouver :

$$\frac{\sum_{i=1}^2 \sum_{j=1}^2 \left( D_j D_i \Phi(\mathbf{s}) \cdot \mathbf{N}_\Phi(\mathbf{s}) \right) v_i v_j}{\|\mathbf{N}_\Phi(\mathbf{s})\| \mathbf{v}^\top Q_{\Phi, \mathbf{s}} \mathbf{v}};$$

La double somme du haut est en fait une forme quadratique du vecteur  $\mathbf{v}$ . Ceci se voit en posant

$$\begin{aligned} e_\Phi(\mathbf{s}) &= D_1 D_1 \Phi(\mathbf{s}) \cdot \mathbf{N}_\Phi(\mathbf{s}) / \|\mathbf{N}_\Phi(\mathbf{s})\| \\ f_\Phi(\mathbf{s}) &= D_1 D_2 \Phi(\mathbf{s}) \cdot \mathbf{N}_\Phi(\mathbf{s}) / \|\mathbf{N}_\Phi(\mathbf{s})\| \\ g_\Phi(\mathbf{s}) &= D_2 D_2 \Phi(\mathbf{s}) \cdot \mathbf{N}_\Phi(\mathbf{s}) / \|\mathbf{N}_\Phi(\mathbf{s})\| \end{aligned} \quad \text{et} \quad \mathbb{I}_{\Phi, \mathbf{s}} = \begin{pmatrix} e_\Phi(\mathbf{s}) & f_\Phi(\mathbf{s}) \\ f_\Phi(\mathbf{s}) & g_\Phi(\mathbf{s}) \end{pmatrix}.$$

$\mathbb{I}$  est nommée **seconde forme fondamentale**. Alors la **courbure** de  $\Phi$  en  $\mathbf{s}$  dans la direction  $\mathbf{v}$  est

$$\kappa_{\Phi, \mathbf{s}}(\mathbf{v}) = \frac{\mathbf{v}^\top \mathbb{I}_{\Phi, \mathbf{s}} \mathbf{v}}{\mathbf{v}^\top Q_{\Phi, \mathbf{s}} \mathbf{v}}.$$

Pour trouver les valeurs maximales ou minimales de  $\kappa$ , une petite astuce pourra nous venir en aide : si  $\mathbf{v} = Q^{-1/2} \mathbf{w}$  (la matrice  $Q^{-1/2}$  existe car  $Q$  est symétrique et définie positive, voir section 8.A), alors

$$\kappa(\mathbf{v}) = \frac{\mathbf{w}^\top Q^{-1/2} \mathbb{I} Q^{-1/2} \mathbf{w}}{\mathbf{w}^\top \mathbf{w}}.$$

Le maximum (resp. minimum) correspond au vecteur propre  $\mathbf{w}$  de  $Q^{-1/2} \mathbb{I} Q^{-1/2}$  dont la valeur propre est maximale (resp. minimale). Deux choses sont faciles à dire sur ces valeurs propres (qui sont parfois nommées **courbures principales**). Leur produit, appelée **courbure de Gauß** est

$$K_\Phi(\mathbf{s}) = K := \lambda_1 \lambda_2 = \text{Det}(Q^{-1/2} \mathbb{I} Q^{-1/2}) = (\text{Det } Q)^{-1} \text{Det } \mathbb{I} = \frac{eg - f^2}{EG - F^2}$$

Leur moyenne, appelée **courbure moyenne**, est (en utilisant,  $\text{Tr}(AB) = \text{Tr}(BA)$ )

$$H_\Phi(\mathbf{s}) = H := \frac{1}{2}(\lambda_1 + \lambda_2) = \frac{1}{2}(\text{Tr}(Q^{-1/2} \mathbb{I} Q^{-1/2})) = \frac{1}{2}(\text{Tr}(Q^{-1} \mathbb{I})) = \frac{1}{2} \frac{Ge + gE - 2fF}{EG - F^2}.$$

Un petit peu d'algèbre permet de retrouver  $\lambda_1$  et  $\lambda_2$  à partir de  $H$  et  $K$  :  $\lambda = H \pm \sqrt{H^2 - K}$ .

La courbure de Gauß, ou plutôt son signe, a un sens très facile à voir géométriquement. En effet, si elle est positive, ceci veut dire que  $\kappa(\mathbf{v})$  est aussi toujours positif, autrement dit que toutes les courbes ont une courbure qui va dans le même sens (c'est le cas de la sphère, voir l'exemple 8.5.2 ci-dessous). Par contre, si elle est négative, ceci veut dire que  $\kappa(\mathbf{v})$  prend des valeurs positives et négatives, autrement dit qu'une courbe a une courbure qui va dans un sens, et une autre dans l'autre sens (c'est le cas de la "selle de cheval", le graphe de la fonction  $F(x, y) = xy$ , voir l'exemple 8.5.3).

**Exemple 8.5.2.** Comme à l'exemple 8.4.1, soit  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  le paramétrage de la sphère défini par  $\Phi(s, t) = \begin{pmatrix} \cos s \sin t \\ \sin s \sin t \\ \cos t \end{pmatrix}$ . Alors

$$D_1\Phi(s, t) = \begin{pmatrix} -\sin s \sin t \\ \cos s \sin t \\ 0 \end{pmatrix}, \quad D_2\Phi(s, t) = \begin{pmatrix} \cos s \cos t \\ \sin s \cos t \\ -\sin t \end{pmatrix}, \quad \text{et } \mathbf{N}_\Phi(s, t) = \begin{pmatrix} -\cos s \sin^2 t \\ -\sin s \sin^2 t \\ -\sin t \cos t \end{pmatrix}.$$

Ce qui donne

$$E_\Phi(s, t) = \sin^2 t, \quad F_\Phi(s, t) = 0, \quad \text{et } G_\Phi(s, t) = 1.$$

Ensuite,

$$D_1D_1\Phi(s, t) = \begin{pmatrix} -\cos s \sin t \\ -\sin s \sin t \\ 0 \end{pmatrix}, \quad D_1D_2\Phi(s, t) = \begin{pmatrix} -\sin s \cos t \\ \cos s \cos t \\ 0 \end{pmatrix}, \quad \text{et } D_2D_2\Phi(s, t) = \begin{pmatrix} -\cos s \sin t \\ -\sin s \sin t \\ -\cos t \end{pmatrix}.$$

Puisque  $\|\mathbf{N}_\Phi(s, t)\|^2 = \sin^2 t$  et  $\sin t \geq 0$  car  $t \in [0, \pi]$ ,

$$e_\Phi(s, t) = \sin^2 t, \quad f_\Phi(s, t) = 0, \quad \text{et } g_\Phi(s, t) = 1.$$

[C'est une coïncidence rare que les deux formes fondamentales soient égales.] Un calcul rapide donne que la courbure de Gauß et la courbure moyenne sont toutes deux égales à 1. En fait, la courbure en un point quelconque et une direction quelconque est

$$\kappa_{\Phi, s}(\mathbf{v}) = 1.$$

Autrement dit, elle est la même partout (ce qui est géométriquement attendu).



**Exemple 8.5.3.** Le caténoïde (voir exemple 8.0.2) possède le paramétrage  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  défini par

$$\Phi(s, t) = \begin{pmatrix} \text{ch } s \cos t \\ \text{ch } s \sin t \\ s \end{pmatrix}.$$

Un premier calcul donne

$$D_1\Phi(s, t) = \begin{pmatrix} \text{sh } s \cos t \\ \text{sh } s \sin t \\ 1 \end{pmatrix}, \quad D_2\Phi(s, t) = \begin{pmatrix} -\text{ch } s \sin t \\ \text{ch } s \cos t \\ 0 \end{pmatrix}, \quad \text{et } \mathbf{N}_\Phi(s, t) = \begin{pmatrix} \text{ch } s \cos t \\ \text{ch } s \sin t \\ \text{ch } s \text{sh } s \end{pmatrix}.$$

D'où la première forme fondamentale :

$$E_\Phi(s, t) = \text{ch}^2 s, \quad F_\Phi(s, t) = 0, \quad \text{et } G_\Phi(s, t) = \text{ch}^2 s.$$

Ensuite,

$$D_1D_1\Phi(s, t) = \begin{pmatrix} \text{ch } s \cos t \\ \text{ch } s \sin t \\ 0 \end{pmatrix}, \quad D_1D_2\Phi(s, t) = \begin{pmatrix} -\text{sh } s \sin t \\ \text{sh } s \cos t \\ 0 \end{pmatrix}, \quad \text{et } D_2D_2\Phi(s, t) = \begin{pmatrix} -\text{ch } s \cos t \\ -\text{ch } s \sin t \\ 0 \end{pmatrix}.$$

Puisque  $\|\mathbf{N}_\Phi(s, t)\|^2 = \text{ch}^2 s(1 + \text{sh}^2 s) = \text{ch}^4 s$  (et  $\text{ch } s \geq 1 > 0$  pour tout  $t \in \mathbb{R}$ ),

$$e_\Phi(s, t) = 1, \quad f_\Phi(s, t) = 0, \quad \text{et } g_\Phi(s, t) = -1.$$

De là, la courbure de Gauß et la courbure moyenne sont :

$$K = \frac{-1}{\text{ch}^4 s} \quad \text{et} \quad H = 0.$$

Les surfaces à courbure moyenne nulles sont très particulières (ce sont les surfaces “minimales”, car elle minimisent l’aire si le bord est fixé). Cette fois-ci la courbure n’est pas uniforme (ce qui est automatique si  $K < 0$ ). Puisque  $H = 0$ , les courbures principales sont  $\lambda_{\pm} = \pm\sqrt{-K} = \pm\text{ch}^{-2}s$ . La courbure à l’image du point  $(s, t)$  et la direction correspondantes sous  $\Phi$  à  $\mathbf{v}$  est

$$\kappa_{\Phi,s}(\mathbf{v}) = \frac{\mathbf{v}^T \mathbb{I} \mathbf{v}}{\mathbf{v}^T Q \mathbf{v}} = \frac{v_1^2 - v_2^2}{\text{ch}^2 s \|\mathbf{v}\|^2}.$$

Les vecteurs qui donnent le maximum et minimum sont évidemment  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  et  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  (le signe de la valeur propre est déterminé par le fait que cette courbure pointe ou non dans la direction du vecteur normal). Autrement dit, les courbes qui correspondent aux courbures les plus extrêmes en un point  $(s, t)$  sont celles de la forme  $\Phi \circ \gamma(s + u, t)$  et  $\Phi \circ \gamma(s, t + u)$  où  $u \in \mathbb{R}$ . ♣

L’autre question mentionnée dans l’introduction est de savoir comment trouver l’équivalent d’une “ligne droite” dans une surface. Cette question est bien sûr reliée. Une personne qui serait née dans un monde en 2 dimension et qui penserait aller en ligne droite (et à vitesse constante) ne sentirait aucune accélération. Il y en aurait forcément une : celle qui fait que cette personne reste dans la surface.

Il serait possible d’utiliser ce qui a été fait au préalable sur le calcul de la courbure, mais il est plus simple d’énoncer ces conditions directement. Il suffit donc de vérifier une condition :

**Définition 8.5.4.** Soit  $\Phi : D \rightarrow \mathbb{R}^m$  un paramétrage régulier. Une courbe  $\gamma : [a, b] \rightarrow D$  est dite **géodésique** si  $(\Phi \circ \gamma)''(t)$  est orthogonal à l’espace tangent en  $\gamma(t)$ ,  $T_{\gamma(t)}^{par} \Phi$ . ★

La condition est équivalente à

$$(8.5.5) \quad \forall j \in \overline{2}, \quad \langle (\Phi \circ \gamma)''(t), D_2 \Phi(\gamma(t)) \rangle = 0$$

Pour voir que cette condition donne automatiquement que la vitesse est constante, il suffit de dériver  $\|(\Phi \circ \gamma)'(t)\|^2$  :

$$\left( \|(\Phi \circ \gamma)'(t)\|^2 \right)' = 2 \langle (\Phi \circ \gamma)'(t), (\Phi \circ \gamma)''(t) \rangle.$$

Mais comme  $(\Phi \circ \gamma)'(t)$  est dans le plan tangent en  $\gamma(t)$ , c’est une combinaison linéaire de  $D_1 \Phi(\gamma(t))$  et  $D_2 \Phi(\gamma(t))$ . Si (8.5.5) tient, le produit scalaire sera alors 0, et la vitesse constante.

La condition des géodésique donne donc les deux équations ( $j \in \overline{2}$ ) :

$$\sum_{k \in \overline{2}} \langle D_k \Phi(\gamma(t)), D_j \Phi(\gamma(t)) \rangle \gamma_k''(t) + \sum_{k, \ell \in \overline{2}} \langle D_k D_\ell \Phi(\gamma(t)), D_j \Phi(\gamma(t)) \rangle \gamma_k'(t) \gamma_\ell'(t) = 0.$$

Une forme vectorielle qui est plus tentante est, en posant  $\Gamma_{k\ell}^j = \langle D_k D_\ell \Phi(\gamma(t)), D_j \Phi(\gamma(t)) \rangle$ ,

$$\sum_{k \in \overline{2}} (Q)_{jk} \gamma_k'' + \sum_{k, \ell \in \overline{2}} \Gamma_{k\ell}^j \gamma_k' \gamma_\ell' = 0.$$

Les  $\Gamma_{k\ell}^j$  sont appelés **symboles de Christoffel**. La somme-produit dans laquelle ils apparaissent ci-haut est typique du calcul tensoriel (une généralisation du calcul matriciel pour des “matrices” qui

sont des tableaux avec trois dimensions). Pour se ramener à quelques chose de purement matriciel, il faudrait prendre  $\Gamma^1$  et  $\Gamma^2$  comme des matrices, puis mettre dans un vecteur les deux réels  $\gamma'^T \Gamma^j \gamma'$  (le premier terme étant de manière plus évidente  $Q\gamma''$ ).

Dans l'utlme chapitre, il sera démontré que les géodésiques sont des “point critiques” pour la longueur. Plus précisément, une géodésique est souvent un minimum (local !) pami les courbes qui relie deux points. Le premier exemple montrera bien que ce ne sont pas toujours des minima.

**Exemple 8.5.6.** (suite de l'exemple 8.5.2.) L'équateur et le méridien sont des géodésiques sur la sphère. Les symboles de Christoffel sont :

$$\begin{array}{lll} \Gamma_{11}^1 = 0 & \Gamma_{12}^1 = \Gamma_{21}^1 = \sin t \cos t & \Gamma_{22}^1 = 0 \\ \Gamma_{11}^2 = -\cos t \sin t & \Gamma_{12}^2 = \Gamma_{21}^2 = 0 & \Gamma_{22}^2 = 0 \end{array}$$

Ce qui donne les deux équations :

$$\begin{array}{l} \gamma_1'' \sin^2 \gamma_2 + 2\gamma_1' \gamma_2' \sin t \cos t = 0 \\ \gamma_2'' - 2\gamma_1'^2 \sin \gamma_2 \cos \gamma_2 = 0 \end{array}$$

Une première solution évidente s'obtient en posant  $\gamma_2' \equiv 0$ . Alors il reste (puisque  $\sin \gamma_2 \neq 0$  (car  $\gamma_2(r) \in ]0, \pi[$ ) :

$$\begin{array}{l} \gamma_1'' = 0 \\ \gamma_1'^2 \cos \gamma_2 = 0. \end{array}$$

D'où  $\cos \gamma_2 = 0$ , d'où  $\gamma_2 \equiv \pi/2$  (et  $\gamma_1(r) = K_1 r + K_2$ ). Ceci est l'équateur.

D'autre part,  $\gamma_1' \equiv 0$ , il reste :  $\gamma_2'' = 0$ . Ce sont les méridiens. ♣

En fait, les méridiens et l'équateur représente la même courbe : l'intersection de la sphère avec un plan qui passe par l'origine. Ces courbes sont appelés des “grand cercles” et forment effectivement le chemin le plus court sur la sphère (comme le lecteur le remarquera si il prend un avion sur une longue distance<sup>5</sup>).

Une manière directe de le voir provient du calcul de la courbure : elle est identiquement égale à 1. D'autre part le vecteur (unitaire) normal en un point  $\mathbf{x}$  de la sphère est  $\mathbf{x}$  lui-même. Autrement dit, pour qu'une coube soit géodésique il faut que son le vecteur normal soit égal à sa position. Ceci est réalisé précisément par un “grand cercle”.

Étant donné deux points sur la sphère, il y a un grand cercle qui les relie<sup>6</sup>. Ce grand cercle aura presque tout le temps une section plus courte que l'autre. La section la plus courte est un minimum (même le minimum), tandis que la plus longue n'est qu'un point d'inflexion<sup>7</sup>.

Voici un exemple moins conclusif, mais sur une surface plus compliquée.

---

5. Les vents dominants font que la trajectoire n'est jamais exactement un grand cercle.

6. Il y a en même exactement un, sauf dans la sitation où les deux points sont à l'opposée sur la sphère : il y a une infinité de méridien entre le pôle nord et le pôle sud.

7. Il est impossible de faire un maximum : il y a toujours une courbe qui oscille encore plus que la précédente et dont la longueur sera encore plus longue.

**Exemple 8.5.7.** (suite de l'exemple 8.5.3.) Cet exemple a pour but de montrer qu'il y a deux géodésiques évidentes sur le caténoïde. Les symboles de Christoffel sont :

$$\begin{array}{lll} \Gamma_{11}^1 = \text{sh } s \text{ ch } s & \Gamma_{12}^1 = \Gamma_{21}^1 = 0 & \Gamma_{22}^1 = -\text{sh } s \text{ ch } s \\ \Gamma_{11}^2 = 0 & \Gamma_{12}^2 = \Gamma_{21}^2 = \text{sh } s \text{ ch } s & \Gamma_{22}^2 = 0 \end{array}$$

Ainsi les équations des géodésiques sont :

$$\begin{array}{l} \gamma_1' \text{ch}^2 \gamma_1 + \gamma_1'^2 \text{sh } \gamma_1 \text{ch } \gamma_1 - \gamma_2'^2 \text{sh } \gamma_1 \text{ch } \gamma_1 = 0 \\ \gamma_2' \text{ch}^2 \gamma_1 + 2\gamma_1' \gamma_2' \text{sh } \gamma_1 \text{ch } \gamma_1 = 0 \end{array}$$

La plus simple solution est de prendre  $\gamma_1' \equiv 0$ , i.e.  $\gamma_1(r) = s_0$ . Puis il reste  $\gamma_2' \equiv 0$ , autrement dit :

$$\gamma(r) = \begin{pmatrix} s_0 \\ K_1 r + K_2 \end{pmatrix}$$

Dans le dessin de l'image à l'exemple 8.0.2 ce sont les cercles horizontaux.

Inversement, en prenant  $\gamma_2' \equiv 0$ , i.e.  $\gamma_2(r) = t_0$ , il reste

$$\gamma_1' \text{ch } \gamma_1 + \gamma_1'^2 \text{sh } \gamma_1 = 0.$$


Il n'est pas vraiment nécessaire de montrer qu'il y a une solution non-triviale à cette équation (un théorème nous l'assure). Il d'autre part automatique que, dans l'image cette correspondra à un bout horizontal de la surface, paramétré, mais pas par la longueur d'arc, par  $\begin{pmatrix} \text{ch } r \cos t_0 \\ \text{ch } r \sin t_0 \\ r \end{pmatrix}$ .

Pour ceux qui veulent en avoir le cœur net, les astuces usuelles permettent de réduire la solution de cette équation à celle de  $v(\gamma)$  (avec  $v = \gamma' \neq 0$ , sinon la solution constante serait obtenue) :

$$v v' \text{ch } \gamma + v^2 \text{sh } \gamma = 0 \quad \Rightarrow \quad \frac{v'}{v} = -\frac{\text{sh } \gamma}{\text{ch } \gamma} \quad \Rightarrow \quad v = K_1 (\text{ch } \gamma)^{-1}.$$

Puis,

$$K_1 r + K_2 = \int \text{ch } \gamma d\gamma = \text{sh } \gamma.$$

Soit  $\gamma_1(r) = \text{argsh}(K_1 r + K_2)$ . Ceci donne une courbe (paramétrée par la longueur d'arc)  $\begin{pmatrix} \sqrt{1+r^2} \cos t_0 \\ \sqrt{1+r^2} \sin t_0 \\ \text{argsh } r \end{pmatrix}$ . 

Malheureusement, contrairement à la sphère sur le caténoïde tous les points ne sont pas identiques. L'exemple ci-dessus ne donne que très peu d'information sur une géodésique quelconque.

EXERCICE 147: *Le tore (de rayons  $r$  et  $R$ ) est une surface qui peut être vue comme l'image du paramétrage  $\Phi : ]-\pi/2, 3\pi/2[ \times ]-\pi, \pi[ \rightarrow \mathbb{R}^3$  défini par  $\Phi(s, t) = \begin{pmatrix} (R+r \cos s) \cos t \\ (R+r \cos s) \sin t \\ r \sin s \end{pmatrix}$ . Montrer que la courbure de Gauss est positive si  $s \in ]-\pi/2, \pi/2[$  et négative si  $s \in ]\pi/2, 3\pi/2[$ .*

EXERCICE 148: (voir exercice 142) Soit  $U \subset \mathbb{R}^2$  un ouvert connexe par arc et  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction de classe  $C^2$ . Soit  $S$  l'ensemble (dans  $\mathbb{R}^3$ ) formé par le graphe de  $D$ . Le paramétrage le plus simple de cette surface est  $\Phi : U \rightarrow S$  défini par  $\Phi(x, y) = \begin{pmatrix} x \\ y \\ F(x, y) \end{pmatrix}$ .

8. Étant donné le domaine, l'image sera un tore privé de deux cercles, mais ce n'est pas bien important.

a. Montrer que

$$K_{\Phi}(x, y) = \frac{D_1 D_1 F(x, y) D_2 D_2 F(x, y) - (D_1 D_2 F(x, y))^2}{(1 + D_1 F(x, y)^2 + D_2 F(x, y)^2)^2}.$$

b. En déduire que si  $F(x, y) = xy$  alors la courbure de Gauss du graphe est partout négative.

c. Toujours si  $F(x, y) = xy$ , calculer les symboles de Christoffel et montrer que  $\gamma(t) = \begin{pmatrix} t \\ 0 \end{pmatrix}$  et  $\sigma(t) = \begin{pmatrix} 0 \\ t \end{pmatrix}$  sont des géodésiques.

d. Montrer que si  $F(x, y) = e^{ax+by}$  (où  $a, b \in \mathbb{R}$ ) alors  $K$  est toujours nulle.

EXERCICE 149: Un des paramétrages de l'hyperboloïde à une nappe (voir l'exemple 8.0.2) est  $\Psi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  défini par

$$\Psi(s, t) = \begin{pmatrix} \operatorname{ch} t \cos s \\ \operatorname{ch} t \sin s \\ \operatorname{sh} t \end{pmatrix}.$$

a. Montrer que, en tout point, la courbure de Gauss est négative.

b. Donner les deux courbures principales et les vecteurs qui y correspondent.

c. L'image de  $\Phi$  peut être décrite comme  $F^{-1}(0)$  où  $F(x, y, z) = x^2 + y^2 - z^2 - 1$ . En déduire que la droite  $u \mapsto \begin{pmatrix} 1 \\ u \\ u \end{pmatrix}$  est contenue dans l'image du paramétrage, et que la présence d'une droite ne force pas l'annulation de  $K$ .

d. Calculer les symboles de Christoffel pour identifier une géodésique simple.

EXERCICE 150: Une des façons de décrire la surface d'Enneper emploie le paramétrage régulier

suivant :  $\Phi(u, v) = \begin{pmatrix} u - \frac{u^3}{3} + uv^2 \\ v - \frac{v^3}{3} + u^2v \\ u^2 - v^2 \end{pmatrix}$  défini de  $\mathbb{R}^2$  dans  $\mathbb{R}^3$ . (Cette application n'est pas injective.)

a. Montrer que la matrice de la première forme fondamentale de  $\Phi$  en  $\begin{pmatrix} u \\ v \end{pmatrix}$  est

$$Q_{\Phi, (u, v)} = (1 + u^2 + v^2)^2 \operatorname{Id} = \begin{pmatrix} (1+u^2+v^2)^2 & 0 \\ 0 & (1+u^2+v^2)^2 \end{pmatrix},$$

c'ad. montrer que  $E_{\Phi}(u, v) = G_{\Phi}(u, v) = (1 + u^2 + v^2)^2$  et que  $F_{\Phi}(u, v) = 0$ .

b. Montrer que la matrice de la seconde forme fondamentale de  $\Phi$  en  $\begin{pmatrix} u \\ v \end{pmatrix}$  est  $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ , c'ad. montrer que  $e_{\Phi}(u, v) = 2 = -g_{\Phi}(u, v)$  et que  $f_{\Phi}(u, v) = 0$ .

c. En déduire que la courbure moyenne est partout 0 et la courbure de Gauss est partout négative.

d. La courbe paramétrée  $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$  définie par  $\gamma(t) = \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix}$  décrit un cercle de longueur  $2\pi r$ . Quelle est la longueur de la courbe  $\sigma = \Phi \circ \gamma$ ? [Indice : se servir de la première forme fondamentale.]

e. Soit  $B_R(\mathbf{0})$  la boule de rayon  $R$  centrée en l'origine  $\mathbf{0}$  dans  $\mathbb{R}^2$ . Soit  $\Psi = \Phi|_{B_R(\mathbf{0})}$  le paramétrage ci-dessus mais restreint à cette boule. Quelle est l'aire de l'image de  $\Psi$ ? [Indice : utiliser un changement de variable en coordonnées polaires.]

EXERCICE 151: (voir exercice 147 pour éviter de refaire les calculs.) Le tore (de rayons  $r$  et  $R$ ) est une surface qui peut être vue comme l'image du paramétrage  $\Phi: ]-\pi/2, 3\pi/2[ \times ]-\pi, \pi[ \rightarrow \mathbb{R}^3$  défini par  $\Phi(s, t) = \begin{pmatrix} (R+r \cos s) \cos t \\ (R+r \cos s) \sin t \\ r \sin s \end{pmatrix}$ . Montrer que  $\gamma(r, t_0)$  est une géodésique ( $\forall t_0$ ) et que  $\gamma(s_0, r)$  est une géodésique seulement si  $s = 0$  ou  $\pi$ .



EXERCICE 152: Calculer les géodésiques du cylindres pour le paramétrage  $\Phi : \mathbb{R} \times ]-\pi, \pi[ \rightarrow \mathbb{R}^3$  donné par  $\Phi(s, t) = \begin{pmatrix} \cos t \\ \sin t \\ s \end{pmatrix}$ .

## 8.A Le calcul fonctionnel pour les matrices symétriques

Une des conséquences intéressante du théorème spectral est que si  $M$  est une matrice symétrique alors  $\Lambda = Q^T M Q$  où  $Q$  est la matrice dont la  $i^{\text{ème}}$  colonne est le  $i^{\text{ème}}$  vecteur propre (noté  $\mathbf{q}_{(i)}$ ) et  $\Lambda$  la matrice diagonale où la  $i^{\text{ème}}$  entrée diagonale est la  $i^{\text{ème}}$  valeur propre. Ceci se montre par un petit calcul : le coefficient  $ij$  de  $Q^T M Q$  est

$$\mathbf{e}_{(j)}^T Q^T M Q \mathbf{e}_{(i)} = (Q \mathbf{e}_{(j)})^T M (Q \mathbf{e}_{(i)}) = \mathbf{q}_{(j)}^T M \mathbf{q}_{(i)} = \lambda_i \mathbf{q}_{(j)} \cdot \mathbf{q}_{(i)}.$$

Or comme les vecteur propres forment une base orthonormée, ce coefficient est 0 si  $i \neq j$  et  $\lambda_i$  si  $i = j$ . De plus,  $Q^T Q = \text{Id} = Q Q^T$ , d'où

$$M = Q \Lambda Q^T.$$

### Lemme 8.A.1

Si une matrice  $M$  est symétrique, alors pour tout entier impair  $k$  il existe une matrice  $M^{1/k}$  qui satisfait  $(M^{1/k})^k = M$ .

Si de plus,  $M$  est semi-définie positive, alors  $\forall r, s \in \mathbb{R}_{\geq 0}$  il existe des matrices  $M^r$  et  $M^s$  telles que  $M^r M^s = M^{r+s}$ , avec  $M^1 = M$  et  $M^0 := \text{Id}$ .

Finalement, si  $M$  est définie positive, alors  $\forall r, s \in \mathbb{R}_{\geq 0}$  il existe des matrices  $M^r$  et  $M^s$  telles que  $M^r M^s = M^{r+s}$ , avec  $M^1 = M$  et  $M^0 := \text{Id}$ .

DÉMONSTRATION: Il suffit essentiellement de remarquer que le produit de matrices diagonales  $D_1$  et  $D_2$  est une matrice diagonale  $D$  dont l'entrée sur la  $i^{\text{ème}}$  ligne et  $i^{\text{ème}}$  colonne est le produit des entrées diagonales correspondantes :  $(D)_{ii} = (D_1)_{ii}(D_2)_{ii}$ .

Ainsi, la matrice  $\Lambda$  possède une matrice  $\Lambda^r$  "évidente" :

$$\Lambda^r = \begin{pmatrix} \lambda_1^r & 0 & \dots & 0 & 0 \\ 0 & \lambda_2^r & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1}^r & 0 \\ 0 & 0 & \dots & 0 & \lambda_n^r \end{pmatrix}$$

Selon que les  $\lambda_i$  sont quelconques ( $M$  symétrique), positifs ou nuls ( $M$  semi-définie positive), strictement positifs ( $M$  définie positive) les différentes puissances (fraction à dénominateur impair, réel positif, réel quelconque) sont définies. Puis, si

$$M^r = Q \Lambda^r Q^T,$$

un petit calcul montre que le comportement est celui attendu :

$$M^r M^s = Q \Lambda^r Q^T Q \Lambda^s Q^T = Q \Lambda^r \text{Id} \Lambda^s Q^T = Q \Lambda^r \Lambda^s Q^T = Q \Lambda^{r+s} Q^T = M^{r+s}.$$

Aussi,  $M^0 = \text{Id}$  et  $M^1 = M$ . ■

De plus,  $\text{Det}(M^r) = (\text{Det} M)^r$ , ainsi, lorsque  $r \neq 0$ ,  $M^r$  est inversible (*i.e.* son déterminant est non-nul) exactement lorsque  $M$  l'est.

L'astuce précédente peut-être généralisée :

**Lemme 8.A.2**

Soit  $f$  et  $g$  des fonctions qui admettent des séries de Taylor qui convergent pour tout  $x \in [a, b]$ . Soit  $M$  une matrice symétrique dont les valeurs propres sont contenues dans l'intervalle  $[a, b]$ . Alors il existe des matrices  $f(M)$  et  $g(M)$  avec  $f(g(M)) = f \circ g(M)$ . Toutes les relations qui n'impliquent que  $f$  et un scalaire<sup>9</sup> sera aussi vraie en remplaçant le scalaire par une matrice.

DÉMONSTRATION: Soit  $f(x) = \sum_{i \geq 0} c_i x^i$ , alors

$$f(M) = \sum_{i \geq 0} c_i M^i = \sum_{i \geq 0} c_i Q \Lambda^i Q^T = Q \left( \sum_{i \geq 0} c_i \Lambda^i \right) Q^T = Q f(\Lambda) Q^T.$$

Ainsi, il est possible de se ramener aux propriétés de  $f$  sur les réels. ■

Comme l'exponentielle est définie par une série de Taylor qui converge partout, il est possible de parler de  $e^M$ ... Des relations surprenantes (et utiles) comme

$$\text{Det } e^M = \prod_{i=1}^n e^{\lambda_i} = e^{\sum_{i=1}^n \lambda_i} = e^{\text{Tr} M}$$

apparaissent. En particulier  $\text{Det } e^M \neq 0$ , et cette matrice est toujours inversible. Même, puisque  $e^{-x} e^x = 1$ ,  $e^{-M} e^M = \text{Id}$ . Comme les valeurs propres de  $e^M$  sont toutes  $> 0$ ,  $(e^M)^r$  existe pour tout réel, et  $(e^M)^r = e^{rM}$ .

Comme  $\ln(1+x)$  a une série de Taylor qui converge pour  $x \in ]-1, 1[$ , si  $M$  est une matrice dont les valeurs propres sont dans cet intervalle,  $\ln(\text{Id} + M)$  est une matrice qui satisfait  $e^{\ln(\text{Id} + M)} = \text{Id} + M$ .

Il est aussi possible de regarder  $e^{\iota M}$  (où  $\iota = \sqrt{-1}$  : c'est une matrice (à coefficients complexes) dont toutes les valeurs propres sont de norme 1.

Par contre, dès que deux matrices interviennent, les relations ne tiennent plus : la base qui diagonalise  $M$  n'est pas forcément la même que celle qui diagonalise  $M'$  (si c'était le cas,  $M$  et  $M'$  commuteraient).

Pour conclure, il est possible de faire du calcul fonctionnel pour des matrices non symétriques. Par exemple, si  $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , alors  $M^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ . Ainsi, toute série de Taylor en  $M$  convergera (puisque toutes les puissances de  $M$  supérieure à 2 sont nulles). Par exemple, si

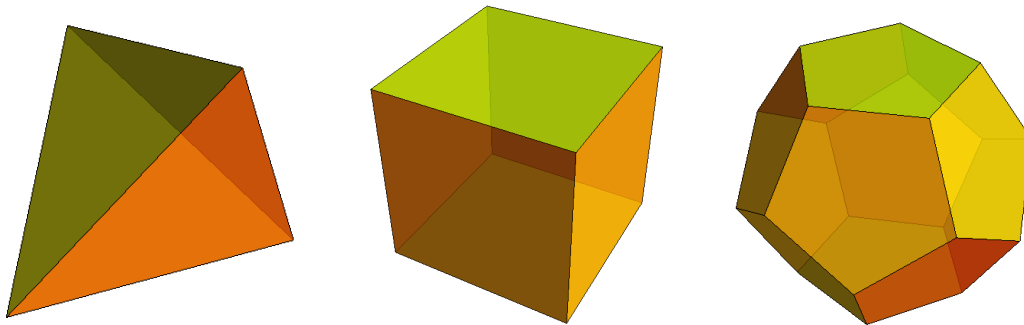
$$N := e^M = \text{Id} + M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{alors} \quad e^{-M} = \text{Id} - M = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = N^{-1}.$$

En poussant un peu plus loin :  $\forall r \in \mathbb{R}$ ,  $e^{rM} = \text{Id} + rM = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} = N^r$ . Pour terminer par une remarque qui va vers les groupes de Lie, pour certains sous-espaces vectoriels de matrices  $V$  bien choisis, il se trouve que l'image par l'exponentielle de  $V$  est un sous-groupe (multiplicatif !) des matrices.

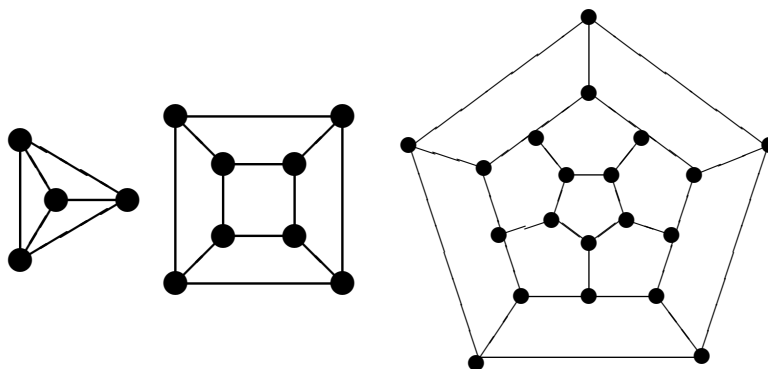
## 8.B Combinatoire et courbure négative

La manière la plus élémentaire de parler de courbure est de regarder un des types les plus anciennement considérés de surface : les polyèdres. Comme souvent, derrière une apparence anodine se cache des sujets assez subtils, et histoire de ne pas trop s'étaler, seul les cas où les faces du polyèdre sont toutes des polygones réguliers identiques sera considéré.

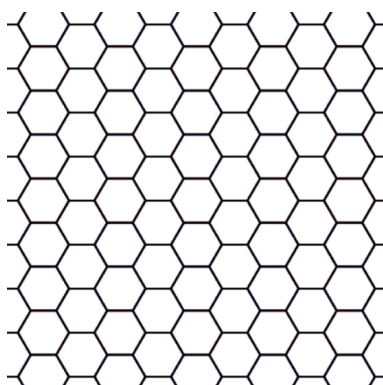
Il n'y a alors que deux données importantes : le nombre de côté au polygone  $f \in \mathbb{Z}_{\geq 3}$  et le nombre de polygone qui se rencontre en un sommet  $d \in \mathbb{Z}_{\geq 3}$ . Par exemple, si  $d = 3$ , les polyèdres construits pour  $f = 3, 4$  ou  $5$  sont le tétraèdre (4 triangles équilatéraux formant une base pyramidale), le cube, et le dodécaèdre :



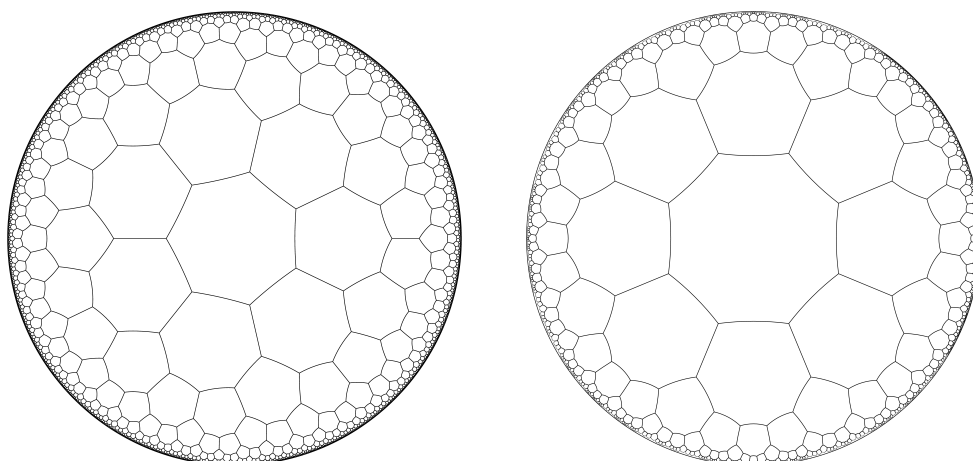
Une manière de les dessiner dans le plan est d'utiliser une projection stéréographique. Plus simplement, en commençant par le polygone à  $f$  côtés puis en continuant avec cette seule règle combinatoire (sans tenir compte des angles) une seule solution s'impose. Ceci donne :



L'angle solide à chaque sommet devient de plus en plus grand lorsque  $f$  augmente. En sorte, la courbure est de moins en moins grande. D'ailleurs, si d'aventure le même procédé était fait avec  $f = 6$ , ce n'est pas un polyèdre qui serait obtenu mais un pavage (comme le lecteur en a certainement déjà vus) :

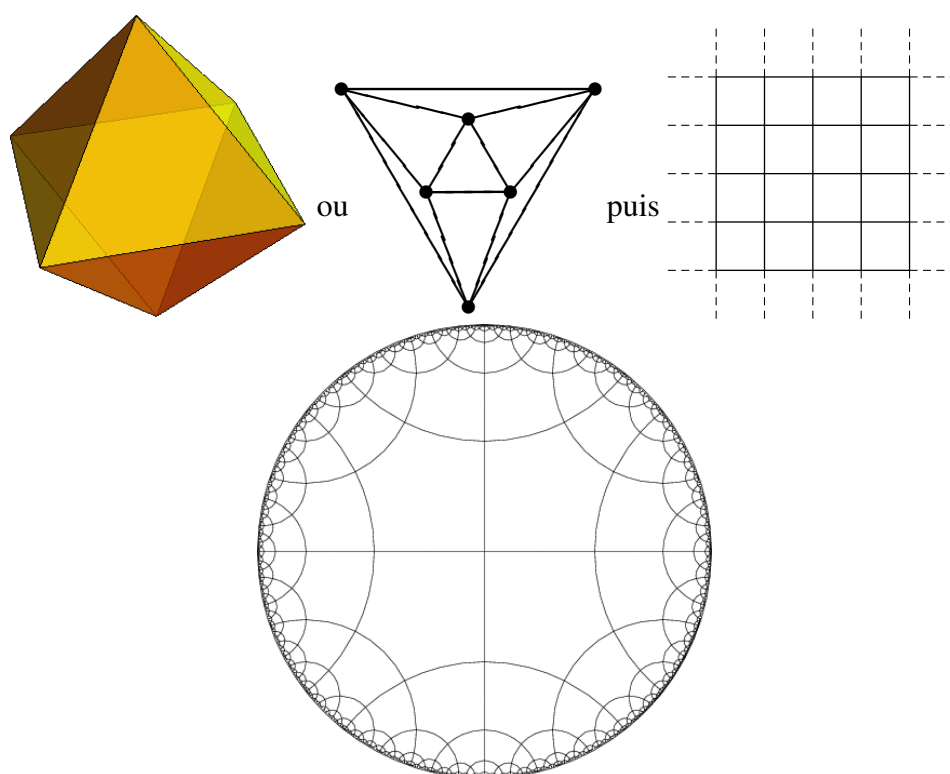


Ainsi, c'est un plan qui est obtenu, et il est raisonnable de dire que la courbure est nulle. Pour poursuivre avec  $f = 7$  ou plus, il faut abandonner de réaliser le "polyèdre" avec une matière rigide. Il est cependant possible de le faire avec du papier et du ruban adhésif. L'objet qui apparaît a alors la propriété qu'un voisinage de tout point a la forme d'une selle de cheval. Voici des représentations déformées pour  $f = 7$  et 8 :

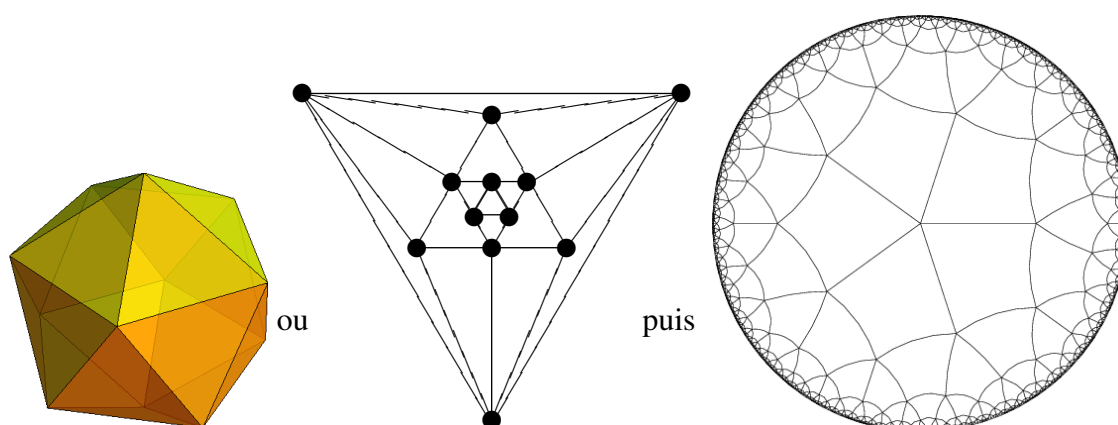


Ces types d'objets sont qualifiés d'hyperboliques. Une raison probable est que la courbure positive, dont les première représentations étaient des ellipsoïdes, fut associée à elliptique. La courbure nulle a deux adjectifs : planaire ou parabolique. Planaire est pour le cas où la courbure est identiquement nulle tandis que parabolique lorsque seule une des deux courbures principales est nulle. Comme les objets qui sont construits ici sont très réguliers (toutes les directions sont similaires) le "parabolique" n'apparaît pas. Finalement, les premiers objets à courbure négatives étaient des hyperboloïdes, d'où le fait que l'adjectif hyperbolique est resté associé à la courbure négative.

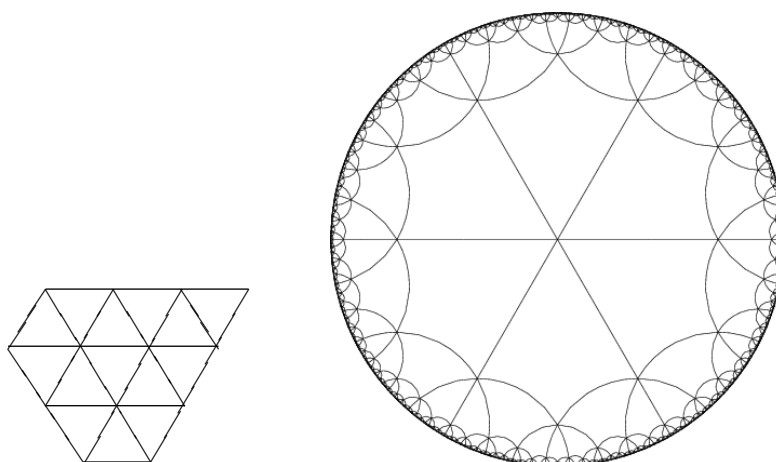
Par souci de complétude, voici la série pour  $d = 4$ . Elle commence par l'octaèdre ( $f = 3$ , deux pyramides à base carrés dont les triangles sont équilatéraux qui sont recollées ensemble par leur base carré), puis le pavage du plan par des carrés (courbure nulle), et les pentagones donnent un objet hyperbolique :



Pour  $d = 5$ , le passage de courbure positive (l'icosaèdre  $f = 3$ ) à la courbure négative ( $f = 4$ ) se fait d'un seul coup :



Pour  $d = 6$ , même avec  $f = 3$  c'est un pavage du plan qui est obtenu puis un objet hyperbolique :



En fait, une idée tentante serait de voir un polyèdre comme un pavage de la sphère. La question naturelle est alors : qu'est-ce que pave les cas où la courbure est négative ?

Ceci nous ramène à une vieille question à savoir s'il est possible d'enlever un des postulats d'Euclide. Celui qui a longtemps fait débat est :

“Étant donné une droite  $D$  et deux droites  $D_1$  et  $D_2$  qui intersectent  $D$  en un point, les droites  $D_1$  et  $D_2$  se coupent quelque part si et seulement si l'angle de ces deux intersection n'est pas le même <sup>10</sup>.”

Il est assez facile de voir que sur une sphère ce postulat est faux : les droites sont des “grand cercles” et n'importe quels grands cercles se croisent. Autrement dit en géométrie sphérique ou elliptique,  $\Rightarrow$  ne tient pas. Cependant, ce qui a longtemps turlupiné les mathématiciens est de savoir si  $\Leftarrow$  pouvait être déduit des autres postulats d'Euclide.

En construisant avec du papier un des objets hyperboliques ci-dessus (préférentiellement celui avec  $d = 3$  et  $f = 7$ , voir plus loin), il est possible de voir que tous les postulats d'Euclide, à l'exception du  $\Leftarrow$ , tiennent. Pour voir de quoi a l'air une droite, la méthode “agricole” recommandée est de prendre un bout de ruban adhésif et de le coller de sorte qu'il n'y ait aucun pli dans le ruban.

“L'objet” pavé dans le cas hyperbolique s'appelle le plan hyperbolique. Un premier avertissement pour les modèles en papier est que le plan hyperbolique ne peut pas être réalisé comme une surface dans  $\mathbb{R}^3$ .

Pour comprendre ce qui fait qu'un objet sphérique, planaire ou hyperbolique est obtenu, le plus simple est de calculer la somme des angles à un sommet. L'angle  $\alpha$  à l'intérieur d'un polygone régulier à  $f$  côtés est donné par  $f(\pi - \alpha) = 2\pi$  ou  $\alpha = (1 - 2/f)\pi$ . Si  $d$  tels polygones se rencontrent en un point, un “cercle” autour de ce point aura pour longueur la somme des angles :  $d\alpha = d(1 - 2/f)\pi$ .

Lorsque  $d(1 - 2/f)\pi = 2\pi$  ceci détermine un pavage du plan. Si  $d(1 - 2/f)\pi < 2\pi$  c'est un “pavage” de la sphère (un polyèdre). La courbure est d'autant plus grande que ce nombre est grand. Finalement, lorsque  $d(1 - 2/f)\pi > 2\pi$  c'est un pavage du “plan hyperbolique”.

Cette dernière caractérisation est typique de la courbure, comme le lecteur le découvrira dans un cours de géométrie différentielle.

10. *i.e.* un angle droit

## Chapitre 9

# Champs de vecteurs

*“Il y a un chant (de vecteurs) endormi en toute chose.”*

*“Le chant de vecteur se reconnaît à son rot.”*

-EINFACH ROT

En physique, les forces qui proviennent d’un potentiel sont particulièrement faciles à traiter entre autres parce que le calcul du travail ne nécessite pas de faire d’intégrale. D’une certaine façon, l’information de la force est déjà bien contenue dans le potentiel. Ces champs de vecteurs portent un nom particulier.

**Définition 9.0.1.** Soit  $U \subset \mathbb{R}^m$  un ouvert. Un champ de vecteurs  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  est dit conservatif si il existe une fonction (dite son potentiel)  $G : U \rightarrow \mathbb{R}$  telle que  $\mathbf{F} = \nabla G$ . ★

La démonstration du résultat classique sur le travail est une conséquence directe du théorème fondamental du calcul intégral.

### Proposition 9.0.2

Soit  $G : U \rightarrow \mathbb{R}$  une fonction de classe  $C^1$  et soit  $\gamma : [a, b] \rightarrow \mathbb{R}^m$  telle que  $\text{Im } \gamma \subset U$ . Alors

$$\int_{\gamma} \nabla G \cdot d\mathbf{s} = G(\gamma(b)) - G(\gamma(a)).$$

DÉMONSTRATION: Par la règle de dérivation en chaîne 3.4.1 et la définition du travail 4.5.3,

$$\int_{\gamma} \nabla G \cdot d\mathbf{s} = \int_a^b \nabla G(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \frac{d}{dt} (G(\gamma(t))) dt = G(\gamma(b)) - G(\gamma(a)). \quad \blacksquare$$

Les théorèmes de Green, Stokes et de la divergence sont des versions en plus grande dimension de cette simple observation.

## 9.1 Divergence, rotationnel et laplacien

Comme d’habitude il y a une manière plus générale d’écrire les opérations qui interviennent ici. Cependant, elle est un peu lourde et demande de maîtriser des concepts plus avancés. Il faudra ici

faire deux définitions du rotationnel selon la dimension (2 ou 3) de l'espace considéré. Tout d'abord en dimension 2,

**Définition 9.1.1.** Soit  $U \subset \mathbb{R}^2$ . Soit  $\mathbf{F} : U \rightarrow \mathbb{R}^2$  un champ de vecteurs de classe  $\mathcal{C}^1$ . Alors le rotationnel de  $\mathbf{F}$ , noté  $\nabla \wedge \mathbf{F}$  ou  $\text{rot} \mathbf{F}$ , est la fonction  $U \rightarrow \mathbb{R}$  définie par

$$\nabla \wedge \mathbf{F}(x, y) = D_1 F_2(x, y) - D_2 F_1(x, y) = \frac{\partial F_2}{\partial x}(x, y) - \frac{\partial F_1}{\partial y}(x, y). \quad \diamond$$

★

En dimension 3, ceci devient

**Définition 9.1.2.** Soit  $U \subset \mathbb{R}^3$  et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^3$  un champ de vecteurs de classe  $\mathcal{C}^1$ . Alors le rotationnel de  $\mathbf{F}$ , noté  $\nabla \wedge \mathbf{F}$  ou  $\text{rot} \mathbf{F}$ , est le champ de vecteurs  $U \rightarrow \mathbb{R}^3$  défini par

$$\nabla \wedge \mathbf{F}(\mathbf{x}) = \text{Det} \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1(\mathbf{x}) & F_2(\mathbf{x}) & F_3(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} D_2 F_3(\mathbf{x}) - D_3 F_2(\mathbf{x}) \\ D_3 F_1(\mathbf{x}) - D_1 F_3(\mathbf{x}) \\ D_1 F_2(\mathbf{x}) - D_2 F_1(\mathbf{x}) \end{pmatrix} \quad \diamond$$

★

Il est très important de remarquer que la définition en dimension 3 recouvre essentiellement celle en dimension 2. En effet, étant donné un champ de vecteurs  $\mathbf{F} : U \rightarrow \mathbb{R}^2$ , il est possible de l'étendre en un champ de vecteurs  $\widehat{\mathbf{F}} : U \times \mathbb{R} \rightarrow \mathbb{R}^3$  en posant  $\widehat{\mathbf{F}}(x_1, x_2, x_3) = \begin{pmatrix} F_1(x_1, x_2) \\ F_2(x_1, x_2) \\ 0 \end{pmatrix}$ . Il apparaît alors que  $\nabla \wedge \widehat{\mathbf{F}}(x_1, x_2, x_3) = \begin{pmatrix} 0 \\ 0 \\ \text{rot} \mathbf{F}(x_1, x_2) \end{pmatrix}$ .

De la même manière qu'il est possible de parler du  $\wedge$  dans  $\mathbb{R}^m$  où  $m \geq 4$ , il est possible de parler de rotationnel (et de divergence) en dimension supérieure. Ceci est relégué à un cours de géométrie différentielle.

Le rotationnel mesure normalement à quel point le champ de vecteurs "tourne". Plus précisément, si le champ de vecteurs est interprété comme, par exemple, la vitesse du vent (à une altitude fixée) dans une région, alors les tourbillons et tornades seraient des points avec un grand (en valeur absolue) rotationnel tandis que sur une plage avec une brise de mer constante le rotationnel serait nul. Le signe du rotationnel ne fait qu'indiquer le sens de la rotation.

Une autre caractéristique importante d'un champ de vecteurs est sa tendance à se "dispenser".

**Définition 9.1.3.** Soit  $U \subset \mathbb{R}^m$  et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  un champ de vecteurs de classe  $\mathcal{C}^1$ . La divergence de  $\mathbf{F}$ , notée  $\nabla \cdot \mathbf{F}$  ou  $\text{div} \mathbf{F}$ , est la fonction  $U \rightarrow \mathbb{R}$  définie par

$$\nabla \cdot \mathbf{F}(\mathbf{x}) = \sum_{i \in \overline{m}} \frac{\partial F_i}{\partial x_i}(\mathbf{x}). \quad \diamond$$

★

Toujours dans l'exemple du vent ci-dessus, une explosion donnerait une région où le champ de vecteurs aurait une divergence positive forte tandis qu'une implosion donnerait une région où le champ de vecteurs aurait une divergence fortement négative.



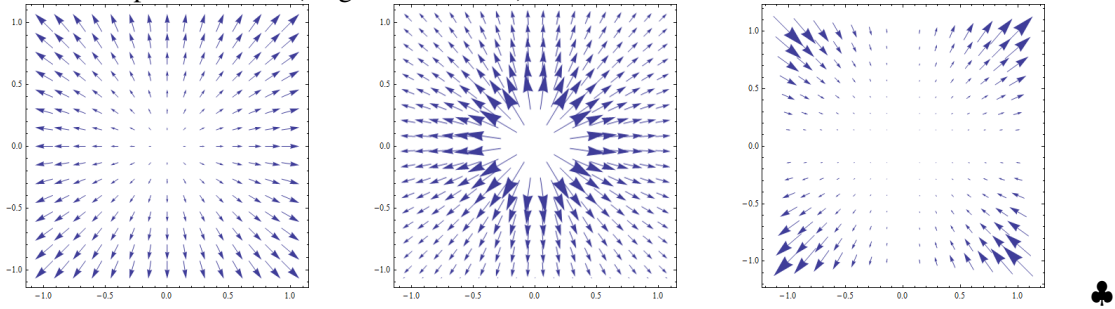
**Exemple 9.1.4.** Soit  $U \subset \mathbb{R}^2$  un ouvert et  $G : U \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ . Soit  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  le champ de vecteurs défini par  $\mathbf{F}(x, y) = \begin{pmatrix} xG(x, y) \\ yG(x, y) \end{pmatrix}$ . Alors

$$\begin{aligned} \operatorname{div} \mathbf{F}(x, y) &= 2G(x, y) + \begin{pmatrix} x \\ y \end{pmatrix} \cdot \nabla G(x, y) \\ \operatorname{rot} \mathbf{F}(x, y) &= -\begin{pmatrix} -y \\ x \end{pmatrix} \cdot \nabla G(x, y) \end{aligned}$$

En particulier, pour  $U = \mathbb{R}^2$  sauf dans le second cas où  $U = \mathbb{R}^2 \setminus \{0\}$ ,

- si  $G(x, y) = 1$  alors  $\operatorname{div} \mathbf{F}(x, y) = 2$  et  $\operatorname{rot} \mathbf{F}(x, y) = 0$  ;
- si  $G(x, y) = (x^2 + y^2)^{-1}$  alors  $\operatorname{div} \mathbf{F}(x, y) = 0$  et  $\operatorname{rot} \mathbf{F}(x, y) = 0$  ;
- si  $G(x, y) = xy$  alors  $\operatorname{div} \mathbf{F}(x, y) = 4xy$  et  $\operatorname{rot} \mathbf{F}(x, y) = y^2 - x^2$ .

Ces trois champs de vecteur (de gauche à droite) :



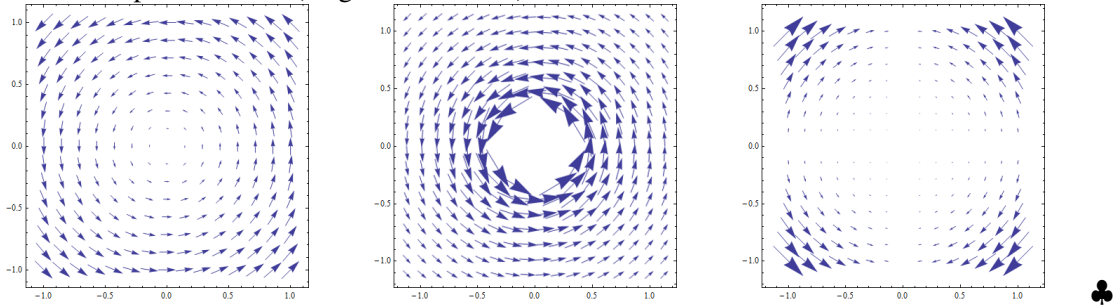
**Exemple 9.1.5.** Soit  $U \subset \mathbb{R}^2$  un ouvert et  $G : U \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ . Soit  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  le champ de vecteurs défini par  $\mathbf{F}(x, y) = \begin{pmatrix} -yG(x, y) \\ xG(x, y) \end{pmatrix}$ . Alors

$$\begin{aligned} \operatorname{div} \mathbf{F}(x, y) &= \begin{pmatrix} -y \\ x \end{pmatrix} \cdot \nabla G(x, y) \\ \operatorname{rot} \mathbf{F}(x, y) &= 2G(x, y) + \begin{pmatrix} x \\ y \end{pmatrix} \cdot \nabla G(x, y) \end{aligned}$$

En particulier, pour  $U = \mathbb{R}^2$  sauf dans le second cas où  $U = \mathbb{R}^2 \setminus \{0\}$ ,

- si  $G(x, y) = 1$  alors  $\operatorname{div} \mathbf{F}(x, y) = 0$  et  $\operatorname{rot} \mathbf{F}(x, y) = 2$  ;
- si  $G(x, y) = (x^2 + y^2)^{-1}$  alors  $\operatorname{div} \mathbf{F}(x, y) = 0$  et  $\operatorname{rot} \mathbf{F}(x, y) = 0$  ;
- si  $G(x, y) = xy$  alors  $\operatorname{div} \mathbf{F}(x, y) = x^2 - y^2$  et  $\operatorname{rot} \mathbf{F}(x, y) = 4xy$ .

Ces trois champs de vecteur (de gauche à droite) :



Un résultat important sur ces opérateurs est que la combinaison de certains d'entre eux est nulle.

### Proposition 9.1.6

Soit  $U \subset \mathbb{R}^3$  et  $\mathbf{F} : U \rightarrow \mathbb{R}^3$  un champ de vecteurs  $\mathcal{C}^2$ . Alors  $\forall \mathbf{x} \in U, \nabla \cdot (\nabla \wedge \mathbf{F})(\mathbf{x}) = \operatorname{div} \operatorname{rot} \mathbf{F}(\mathbf{x}) =$

0.

Soit  $m = 2$  ou  $3$ , soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction  $\mathcal{C}^2$ . Alors  $\forall \mathbf{x} \in U, \nabla \wedge (\nabla F)(\mathbf{x}) = \text{rot grad} F = 0$ .

DÉMONSTRATION: Cette démonstration se résume à un calcul direct... Il ne sera fait que pour  $m = 3$ .

$$\nabla \wedge (\nabla F)(\mathbf{x}) = \text{Det} \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ \frac{\partial F}{\partial x_1}(\mathbf{x}) & \frac{\partial F}{\partial x_2}(\mathbf{x}) & \frac{\partial F}{\partial x_3}(\mathbf{x}) \end{pmatrix}$$

étant donnée que  $\mathbf{F}$  est de classe  $\mathcal{C}^2$ , l'ordre des dérivées est sans importance, et elles se trouvent toutes présentes avec un signe une fois  $+$  et une fois  $-$ .

$$\nabla \cdot (\nabla \wedge \mathbf{F})(\mathbf{x}) = \text{Det} \begin{pmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1(\mathbf{x}) & F_2(\mathbf{x}) & F_3(\mathbf{x}) \end{pmatrix} = 0.$$

En effet, comme l'ordre des dérivées secondes peut s'inverser, le même phénomène se produit. ■

L'exercice suivant montre l'importance de l'hypothèse sur la régularité  $\mathcal{C}^2$  dans le résultat précédent.

EXERCICE 153: À une fonction  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$ , soit  $\mathbf{G} = \nabla F$  le champ de vecteurs (de classe  $\mathcal{C}^0$ ) associé. Son rotationnel,  $\text{rot} \mathbf{G}$ , est défini ponctuellement par les dérivées partielles  $\text{rot} \mathbf{G}(x, y) = \frac{\partial G_2}{\partial x}(x, y) - \frac{\partial G_1}{\partial y}(x, y)$  si elles existent. Montrer qu'il est possible que (dans ce sens)  $\text{rot} \nabla F$  existe en tout point mais ne soit pas nul partout.

L'exercice suivant sert à montrer que le rotationnel et la divergence sont des concepts qui ne dépendent pas du repère choisi, mais que l'hypothèse de régularité est alors importante.

EXERCICE 154: Soit  $\theta \in [0, 2\pi]$ , soit  $\mathbf{u} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$  et  $\mathbf{v} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$  deux vecteurs unitaires, soit  $U$  un ouvert contenant l'origine et soit  $\mathbf{F} : U \rightarrow \mathbb{R}^2$  un champ de vecteurs.

- Montrer que si  $\mathbf{F}$  est de classe  $\mathcal{C}^1$ , alors,  $\forall \theta \in [0, 2\pi]$ , le rotationnel de  $\mathbf{F}$  s'écrit aussi  $D_{\mathbf{u}}(\mathbf{v} \cdot \mathbf{F}(x, y)) - D_{\mathbf{v}}(\mathbf{u} \cdot \mathbf{F}(x, y))$ .
- Trouver une fonction qui n'est pas dérivable en un point, dont toutes les dérivées directionnelles existent et telle que la quantité  $D_{\mathbf{u}}(\mathbf{v} \cdot \mathbf{F}(x, y)) - D_{\mathbf{v}}(\mathbf{u} \cdot \mathbf{F}(x, y))$  dépend de  $\theta$ .

**Remarque 9.1.7.** La proposition 9.1.6 donne l'annulation de l'application successive de certains des opérateurs ci-dessus. Cependant, il est possible de prendre une paire d'opérateurs dont l'application successive n'est pas toujours nulle. En effet, soit  $U \subset \mathbb{R}^m$  un ouvert et  $F : U \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^2$ , alors  $\text{div grad} F(\mathbf{x}) = \nabla \cdot \nabla F(\mathbf{x}) = \sum_{i \in \overline{m}} \frac{\partial^2 F_i}{\partial x_i^2}(\mathbf{x})$  est le laplacien de  $F$ . Les fonctions pour lesquelles cette composition est nulle sont dites les fonctions harmoniques. Il existe beaucoup de fonctions qui ne sont pas harmoniques.

Il serait aussi possible de se demander, lorsque  $m = 2$ , si un champ de vecteurs qui est de rotationnel et de divergence nul est forcément nul. Intuitivement la réponse est clairement non : dans l'interprétation de la vitesse du vent, il est clair qu'un vent de force et de direction constante a

un rotationnel nul et une divergence nulle sans pour autant être trivial. Il est un peu moins évident de trouver d'autres champs de vecteurs ayant cette propriété (bien que les exemples 9.1.4 et 9.1.5 en contiennent tous les deux un). De fait, ceux-ci sont nombreux et importants : un tel champ de vecteurs est équivalent à une fonction holomorphe. ♠

EXERCICE 155: Soit  $A$  une matrice  $2 \times 2$ , et soit  $\mathbf{F}(x, y) = A \begin{pmatrix} x \\ y \end{pmatrix}$  le champ de vecteurs (une application linéaire) associée. Donner  $\text{rot } \mathbf{F}$  et  $\text{div } \mathbf{F}$  en fonction des coefficients de la matrice  $A$ . Montrer que si ils sont tous deux nuls,  $A^2 = c \text{Id}$  où  $c \in \mathbb{R}_{\geq 0}$  et  $\text{Id}$  est la matrice identité.

EXERCICE 156: Soit  $i$  le symbole dénotant une des racines carrées de  $-1$ . Soit  $\mathbb{C}$  l'ensemble des nombres complexes, c'ad.  $z \in \mathbb{C}$  s'écrit formellement  $x + iy$  où  $x, y \in \mathbb{R}$ . Toute fonction  $f : U \rightarrow \mathbb{C}$  (où  $U$  est un ouvert de  $\mathbb{C}$ ) peut s'interpréter comme un champ de vecteurs. En effet, pour  $f$  donné soit  $\mathbf{F}(x, y) = \begin{pmatrix} \Re f(x+iy) \\ -\Im f(x+iy) \end{pmatrix}$  où, si  $a$  et  $b \in \mathbb{R}$ ,  $\Re(a+ib) = a$  et  $\Im(a+ib) = b$ . Montrer que  $\text{rot } \mathbf{F} = \text{div } \mathbf{F} = 0$  si et seulement si  $f$  est holomorphe. [Indice : écrire les équations de Cauchy-Riemann.]

## 9.2 Théorème de Green et de la divergence, cas simple

Le théorème de Green s'appelle parfois aussi le théorème de Green-Riemann.

### Lemme 9.2.1 (Théorème de Green, cas du pavé)

Soit  $P \subset \mathbb{R}^2$  un pavé, soit  $\mathbf{F} : P \rightarrow \mathbb{R}^2$  un champ de vecteurs de classe  $C^1$ , et soit  $\gamma$  une courbe  $C^1$  par morceaux, injective, dont la trace  $\text{Im } \gamma = \partial P$  et qui est compatible avec l'orientation canonique (de  $\mathbb{R}^2$ ). Alors

$$\int_P \text{rot } \mathbf{F} d\text{vol} = \int_\gamma \mathbf{F} \cdot d\mathbf{s}.$$

DÉMONSTRATION: Soit  $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  le coin inférieur gauche du pavé et  $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$  le coin supérieur droit (où  $\forall i \in \underline{2}, a_i < b_i$ ). Séparons  $\gamma$  en quatre morceaux (un par côté du pavé)

$$\begin{aligned} \gamma_{(1)} : [a_1, b_1] &\rightarrow \mathbb{R}^2 & \text{où } \gamma_{(1)}(t) &= \begin{pmatrix} t \\ a_2 \end{pmatrix} \\ \gamma_{(2)} : [a_2, b_2] &\rightarrow \mathbb{R}^2 & \text{où } \gamma_{(2)}(t) &= \begin{pmatrix} b_1 \\ t \end{pmatrix} \\ \gamma_{(3)} : [a_1, b_1] &\rightarrow \mathbb{R}^2 & \text{où } \gamma_{(3)}(t) &= \begin{pmatrix} a_1+b_1-t \\ b_2 \end{pmatrix} \\ \gamma_{(4)} : [a_2, b_2] &\rightarrow \mathbb{R}^2 & \text{où } \gamma_{(4)}(t) &= \begin{pmatrix} a_1 \\ a_2+b_2-t \end{pmatrix} \end{aligned}$$

Alors  $\int_\gamma \mathbf{F} \cdot d\mathbf{s} = \sum_{i \in \underline{4}} \int_{\gamma_{(i)}} \mathbf{F} \cdot d\mathbf{s}$ , et

$$\begin{aligned} \int_{\gamma_{(1)}} \mathbf{F} \cdot d\mathbf{s} &= \int_{a_1}^{b_1} F_1(t, a_2) dt \\ \int_{\gamma_{(2)}} \mathbf{F} \cdot d\mathbf{s} &= \int_{a_2}^{b_2} F_2(b_1, t) dt \\ \int_{\gamma_{(3)}} \mathbf{F} \cdot d\mathbf{s} &= \int_{a_1}^{b_1} -F_1(a_1+b_1-t, b_2) dt = \int_{a_1}^{b_1} -F_1(t, b_2) dt \\ \int_{\gamma_{(4)}} \mathbf{F} \cdot d\mathbf{s} &= \int_{a_2}^{b_2} -F_2(a_1, a_2+b_2-t) dt = \int_{a_2}^{b_2} -F_2(a_1, t) dt \end{aligned}$$

D'un autre côté, par le théorème de Fubini,

$$\begin{aligned} \int_P \text{rot } \mathbf{F} &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \frac{\partial F_1}{\partial y}(x, y) dy dx - \int_{a_2}^{b_2} \int_{a_1}^{b_1} \frac{\partial F_2}{\partial x}(x, y) dx dy \\ &= \int_{a_1}^{b_1} (F_1(x, b_2) - F_1(x, a_2)) dx - \int_{a_2}^{b_2} (F_2(b_1, y) - F_2(a_1, y)) dy. \end{aligned}$$

Quant à lui, le théorème de la divergence porte parfois le nom de théorème de Gauss-Ostrogradsky (ou de Gauss tout court). La démonstration requiert de parler de l'orientation du bord d'un volume dans  $\mathbb{R}^3$ . Il est possible de le faire pour une classe satisfaisante de volumes. De manière intuitive le bord du volume est orienté de sorte que son vecteur normal pointe vers l'extérieur du volume.

Par exemple, soit  $P = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$  un pavé (où  $\forall i \in \overline{3}, a_i < b_i$ ). Soit  $P_{i,\pm}$  les six feuilles composant le bord de  $P$  :

$$\begin{aligned} P_{1,+} &= \{b_1\} \times [a_2, b_2] \times [a_3, b_3] & P_{2,+} &= [a_1, b_1] \times \{b_2\} \times [a_3, b_3] & P_{3,+} &= [a_1, b_1] \times [a_2, b_2] \times \{b_3\} \\ P_{1,-} &= \{a_1\} \times [a_2, b_2] \times [a_3, b_3] & P_{2,-} &= [a_1, b_1] \times \{a_2\} \times [a_3, b_3] & P_{3,-} &= [a_1, b_1] \times [a_2, b_2] \times \{a_3\} \end{aligned}$$

Alors, les feuilles sont orientées de sorte que la normale de  $P_{i,\pm}$  pointe dans la direction de  $\pm \mathbf{e}_i$ .

**Lemme 9.2.2 (Théorème de la divergence, cas du pavé)**

Soit  $P \subset \mathbb{R}^3$  un pavé, soit  $\partial P$  son bord avec l'orientation canonique et soit  $\mathbf{F} : P \rightarrow \mathbb{R}^3$  un champ de vecteurs de classe  $C^1$ . Alors

$$\int_P \nabla \cdot \mathbf{F} d\text{vol} = \int_{\partial P} \mathbf{F} \cdot d\mathbf{N}.$$

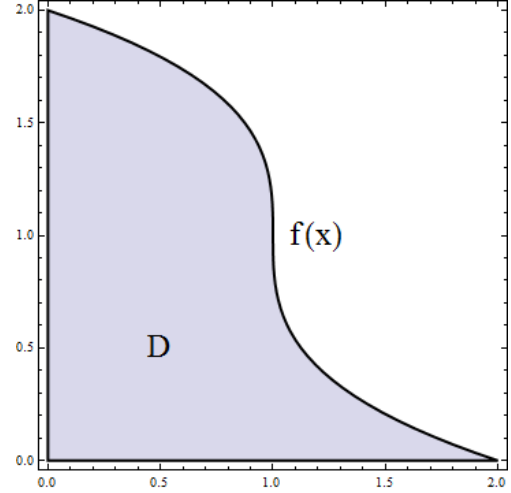
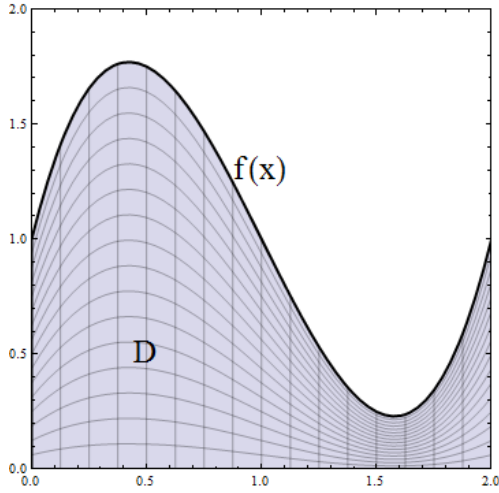
DÉMONSTRATION: Encore une fois, grâce à Fubini

$$\begin{aligned} \int_P (\nabla \cdot \mathbf{F}) d\text{vol} &= \int_{a_2}^{b_2} \int_{a_3}^{b_3} \int_{a_1}^{b_1} \frac{\partial F_1}{\partial x_1} dx_1 dx_3 dx_2 + \int_{a_1}^{b_1} \int_{a_3}^{b_3} \int_{a_2}^{b_2} \frac{\partial F_2}{\partial x_2} dx_2 dx_3 dx_1 \\ &\quad + \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} \frac{\partial F_3}{\partial x_3} dx_3 dx_2 dx_1 \\ &= \int_{a_2}^{b_2} \int_{a_3}^{b_3} \left( F_1(b_1, x_2, x_3) - F_1(a_1, x_2, x_3) \right) dx_3 dx_2 \\ &\quad + \int_{a_1}^{b_1} \int_{a_3}^{b_3} \left( F_2(x_1, b_2, x_3) - F_2(x_1, a_2, x_3) \right) dx_3 dx_1 \\ &\quad + \int_{a_1}^{b_1} \int_{a_2}^{b_2} \left( F_3(x_1, x_2, b_3) - F_3(x_1, x_2, a_3) \right) dx_2 dx_1 \\ &= \sum_{i \in \overline{3}} \int_{P_{i,+}} \mathbf{F} \cdot d\mathbf{N} + \int_{P_{i,-}} \mathbf{F} \cdot d\mathbf{N} \\ &= \int_{\partial P} \mathbf{F} \cdot d\mathbf{N} \quad \blacksquare \end{aligned}$$

Les exercices suivants servent à donner des démonstrations alternatives du théorème de Green, l'une d'entre elles sous des conditions plus faibles.

EXERCICE 157: Soit  $f : [0, 1] \rightarrow \mathbb{R}$  une fonction de classe  $C^2$  telle que  $\forall x \in [0, 1], f(x) > 0$ . Soit  $D$  la région bornée de  $\mathbb{R}^2$  délimitée par les équations  $x = 0$ ,  $x = 1$ ,  $y = 0$  et  $y = f(x)$  (dessin de gauche ci-dessous). Montrer qu'il existe un difféomorphisme  $C^2$  de  $[0, 1] \times [0, 1]$  vers  $D$  dont le jacobien (i.e. le déterminant de la matrice de Jacobi) est positif. En particulier le théorème de Green est vrai pour  $D$ .

Sans refaire tout le calcul, comment étendre ce résultat au cas où  $f(1) = 0$  ?



EXERCICE 158: Voici une manière d'affaiblir les conditions de régularités pour le théorème de Green. Soit  $f : [0, 1] \rightarrow \mathbb{R}$  une fonction continue, strictement décroissante, dérivable sauf en un nombre fini de points, et telle que  $\forall x \in [0, 1[, f(x) > 0$  et  $f(1) = 0$ . Soit  $D$  la région bornée de  $\mathbb{R}^2$  délimitée par les équations  $x = 0$ ,  $y = 0$ , et  $y = f(x)$  (dessin de droite ci-dessus). Pour un champ de vecteurs  $\mathbf{F}$  de classe  $C^1$ , montrer par calcul direct (en utilisant le théorème de Fubini sur ce domaine) que

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{s} = \int_D \text{rot } \mathbf{F} \, d\text{vol}.$$

EXERCICE 159: Vous pouvez supposer connu que les rotations sont des difféomorphismes  $C^\infty$  et le théorème de Jordan. Ce dernier stipule qu'une courbe  $\gamma : [0, L] \rightarrow \mathbb{R}^2$  injective sur l'intervalle ouvert et fermée (i.e.  $\gamma(0) = \gamma(L)$ ) sépare le plan en deux composantes (une bornée et une non-bornée). Soit  $D$  la partie bornée. Supposons de plus que l'ensemble  $\{t \in [0, L] \mid \gamma'_1(t) = 0 \text{ ou } \gamma'_2(t) = 0\}$  est fini. Comment découper l'ensemble  $D$  de façon à montrer le théorème de Green pour  $D$  (utiliser le résultat des deux exercices précédents).

EXERCICE 160: Soit  $P$  un pavé et  $\phi : P \rightarrow \mathbb{R}$  une fonction de classe  $C^2$  et  $\mathbf{F} : P \rightarrow \mathbb{R}^3$  un champ de vecteur de classe  $C^1$ . Les notations  $\text{grad } \phi = \nabla \phi$  et  $\text{div } \mathbf{F} = \nabla \cdot \mathbf{F}$  seront utilisées.

a. En utilisant l'identité  $\text{div}(\phi(\mathbf{x})\mathbf{F}(\mathbf{x})) = (\text{grad } \phi(\mathbf{x})) \cdot \mathbf{F}(\mathbf{x}) + \phi(\mathbf{x}) \text{div } \mathbf{F}(\mathbf{x})$ , montrer que

$$\int_P (\text{grad } \phi(\mathbf{x})) \cdot \mathbf{F}(\mathbf{x}) \, d\text{vol}(\mathbf{x}) = \int_{\partial P} \phi(\mathbf{x}) \mathbf{F}(\mathbf{x}) \cdot d\mathbf{N} - \int_P \phi(\mathbf{x}) \text{div } \mathbf{F}(\mathbf{x}) \, d\text{vol}(\mathbf{x})$$

où le premier terme à droite représente la somme des flux de  $\mathbf{F}$  à travers les différentes surfaces formant le bord du pavé.

b. Supposons que  $\text{div}(\text{grad } \phi(\mathbf{x})) = 0$ , montrer, en posant  $\mathbf{F}(\mathbf{x}) = \text{grad } \phi(\mathbf{x})$ , que

$$\int_P \|\text{grad } \phi(\mathbf{x})\|^2 \, d\text{vol}(\mathbf{x}) = \int_{\partial P} \phi(\mathbf{x}) \text{grad } \phi(\mathbf{x}) \cdot d\mathbf{N}$$

c. Si de plus  $\forall \mathbf{x} \in \partial P$ ,  $\phi(\mathbf{x}) = 0$ , en conclure que  $\forall \mathbf{x} \in P$ ,  $\text{grad } \phi(\mathbf{x}) = 0$ , puis que  $\phi$  est constante et égale à 0.

### 9.3 Paramétrage et théorème de Stokes

Le théorème de Stokes fait généralement référence à un théorème de géométrie différentielle (qui ne sera pas décrit ici). Le théorème dont il sera ici question s'appelle parfois aussi théorème de Kelvin-Stokes ou théorème du rotationnel. Sa démonstration la plus simple est une conséquence du théorème de Green, et permet de terminer la démonstration du théorème de Green pour des domaines plus généraux. Il s'agit de montrer que le théorème de Green (pour le pavé) se comporte bien sous les paramétrages locaux.

#### Théorème 9.3.1 (Théorème de Stokes, cas d'un morceau)

Soit  $P \subset \mathbb{R}^2$  un pavé, et soit  $U \supset P$  un ouvert. Soit  $\Phi : U \rightarrow \mathbb{R}^3$  un paramétrage local **de classe**  $C^2$ ,  $\gamma : [0, L] \rightarrow \Phi(\partial P)$  une courbe dont l'orientation est compatible avec l'orientation induite par  $\Phi$  de  $\partial\Phi(P)$ , soit  $W \subset \mathbb{R}^3$  un ouvert contenant  $\Phi(P)$  et soit  $\mathbf{F} : W \rightarrow \mathbb{R}^3$  un champ de vecteurs de classe  $C^1$ . Alors

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{s} = \int_{\Phi|_P} (\nabla \wedge \mathbf{F}) \cdot d\mathbf{N}$$

DÉMONSTRATION: Dans cette démonstration, les éléments de  $\mathbb{R}^3$  seront écrits  $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ , et ceux de  $\mathbb{R}^2$   $\begin{pmatrix} s_1 \\ s_2 \end{pmatrix}$ . Soit  $\sigma : [0, L] \rightarrow \partial P \subset \mathbb{R}^2$  une courbe  $C^1$  par morceaux donnant l'orientation canonique de  $\partial P$  et telle que  $\gamma = \Phi \circ \sigma$  (elle existe car  $\Phi$  est essentiellement un difféomorphisme et  $\gamma$  est  $C^1$  par morceaux). Alors

$$\begin{aligned} \int_{\gamma} \mathbf{F} \cdot d\mathbf{s} &= \int_0^L \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt && \text{par définition} \\ &= \int_0^L \mathbf{F}(\Phi \circ \sigma(t)) \cdot \left( \sigma'_1(t) D_1 \Phi(\sigma(t)) + \sigma'_2(t) D_2 \Phi(\sigma(t)) \right) dt && \text{car } \gamma'(t) = \sigma'(t) \Phi_*(\gamma(t))^T \\ &= \int_0^L \begin{pmatrix} \mathbf{F}(\Phi \circ \sigma(t)) \cdot D_1 \Phi(\sigma(t)) \\ \mathbf{F}(\Phi \circ \sigma(t)) \cdot D_2 \Phi(\sigma(t)) \end{pmatrix} \cdot \sigma'(t) dt && \text{réécriture.} \\ &= \int_0^L \mathbf{G}(\sigma(t)) \cdot \sigma'(t) dt && \text{voir déf. de } \mathbf{G} \text{ ci-dessous.} \\ &= \int_{\sigma} \mathbf{G} \cdot d\mathbf{s} && \text{par définition.} \end{aligned}$$

où  $\mathbf{G} : P \rightarrow \mathbb{R}^2$  est un nouveau champ de vecteurs (de  $\mathbb{R}^2$ ) défini par

$$\mathbf{G}(s_1, s_2) = \begin{pmatrix} \mathbf{F}(\Phi(s_1, s_2)) \cdot D_1 \Phi(s_1, s_2) \\ \mathbf{F}(\Phi(s_1, s_2)) \cdot D_2 \Phi(s_1, s_2) \end{pmatrix}.$$

L'idée est maintenant de se ramener au théorème de Green pour le pavé  $P$ , *i.e.* de montrer que le côté gauche est aussi le rotationnel de  $\mathbf{G}$  sur le pavé. Il s'agit d'un calcul direct, mais pas tout à fait

simple.

$$\begin{aligned}
 \operatorname{rot} \mathbf{G}(s_1, s_2) &= \frac{\partial}{\partial s_1} G_2(s_1, s_2) - \frac{\partial}{\partial s_2} G_1(s_1, s_2) && \text{par définition} \\
 &= \mathbf{F}(\Phi(s_1, s_2)) \cdot D_1 D_2 \Phi(s_1, s_2) && \text{dérivation du P.S.} \\
 &\quad + \sum_{i \in \bar{3}} \left[ \frac{\partial \Phi_i}{\partial s_1}(s_1, s_2) \frac{\partial \mathbf{F}}{\partial x_i}(\Phi(s_1, s_2)) \cdot D_2 \Phi(s_1, s_2) \right. && \text{dérivation en chaîne} \\
 &\quad \quad \left. - \frac{\partial \Phi_i}{\partial s_2}(s_1, s_2) \frac{\partial \mathbf{F}}{\partial x_i}(\Phi(s_1, s_2)) \cdot D_1 \Phi(s_1, s_2) \right] \\
 &\quad - \mathbf{F}(\Phi(s_1, s_2)) \cdot D_2 D_1 \Phi(s_1, s_2) \\
 &= \sum_{i \in \bar{3}} \sum_{j \in \bar{3}} \left[ D_1 \Phi_i(s_1, s_2) D_2 \Phi_j(s_1, s_2) \frac{\partial F_j}{\partial x_i}(\Phi(s_1, s_2)) \right. && \text{développer le P.S.} \\
 &\quad \quad \left. - D_2 \Phi_i(s_1, s_2) D_1 \Phi_j(s_1, s_2) \frac{\partial F_j}{\partial x_i}(\Phi(s_1, s_2)) \right] \\
 &= \sum_{i \neq j} \left[ \frac{\partial F_j}{\partial x_i}(\Phi(s_1, s_2)) - \frac{\partial F_i}{\partial x_j}(\Phi(s_1, s_2)) \right] D_1 \Phi_i(s_1, s_2) D_2 \Phi_j(s_1, s_2)
 \end{aligned}$$

Maintenant, il faut faire un second calcul (plus court). Avant de s'y lancer, il est bon de jeter un oeil à l'écriture du produit mixte  $(\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}))$ . En effet,  $(\nabla \wedge \mathbf{F}(\Phi(s_1, s_2))) \cdot \mathbf{N}_\Phi(s_1, s_2)$  est égal à

$$\operatorname{Det} \begin{pmatrix} \frac{\partial F_3}{\partial x_2}(\Phi(\dots)) - \frac{\partial F_2}{\partial x_3}(\Phi(\dots)) & \frac{\partial F_1}{\partial x_3}(\Phi(\dots)) - \frac{\partial F_3}{\partial x_1}(\Phi(\dots)) & \frac{\partial F_2}{\partial x_1}(\Phi(\dots)) - \frac{\partial F_1}{\partial x_2}(\Phi(\dots)) \\ D_1 \Phi_1(s_1, s_2) & D_1 \Phi_2(s_1, s_2) & D_1 \Phi_3(s_1, s_2) \\ D_2 \Phi_1(s_1, s_2) & D_2 \Phi_2(s_1, s_2) & D_2 \Phi_3(s_1, s_2) \end{pmatrix}.$$

Or ce déterminant est la même chose que  $\operatorname{rot} \mathbf{G}(s_1, s_2)$ . Mais comme le théorème de Green a été démontré pour un pavé, l'affirmation de ce théorème s'ensuit. ■

Le cadre plus général se fait en découpant les surface en morceaux

### Corollaire 9.3.2 (Théorème de Stokes)

Soit  $S$  une surface orientable. Supposons que  $\partial S$  contient  $k \in \mathbb{Z}_{\geq 0}$  composantes connexes<sup>1</sup>. Pour  $j \in \bar{k}$  soit  $\gamma_j$  des courbes  $C^1$  par morceaux qui sont compatibles avec l'orientation induite sur la  $j^{\text{ème}}$  composante connexe. Soit  $\mathbf{F}$  un champ de vecteurs défini sur un voisinage de  $S$ . Alors,

$$\int_S (\nabla \wedge \mathbf{F}) \cdot d\mathbf{N} = \sum_{j \in \bar{k}} \int_{\gamma_j} \mathbf{F} \cdot d\mathbf{s}. \quad \left[ \text{le terme de droite se note } \int_{\partial S} \mathbf{F} \cdot d\mathbf{s} \right]$$

Le théorème ci-dessus ne représente pas les conditions de régularité optimales. Si  $S$  n'a pas de bord, le terme de droite se lit alors 0.

Pour passer au théorème de Green plus général, il suffit de constater que si  $\mathbf{H} : P \rightarrow \mathbb{R}^2$  est un difféomorphisme (sur son image) alors, il se transforme en un paramétrage local  $\Phi : P \rightarrow \mathbb{R}^3$  défini par  $\Phi(s_1, s_2) = \begin{pmatrix} H_1(s_1, s_2) \\ H_2(s_1, s_2) \\ 0 \end{pmatrix}$ . Ensuite, comme le rotationnel de  $\mathbb{R}^2$  est cohérent avec celui de  $\mathbb{R}^3$  et que la normale de  $\Phi$  pointe précisément dans la bonne direction, le théorème de Green est étendu à tout domaine difféomorphe à un pavé. Au passage, notons que tout pavé de  $\mathbb{R}^m$  est difféomorphe à  $[0, 1]^m$

**Corollaire 9.3.3 (Théorème de Green, cas (presque) général)**

Soit  $D \subset \mathbb{R}^2$  un domaine tel qu'il existe  $k \in \mathbb{Z}_{\geq 1}$  et  $\forall i \in \bar{k}$  des fermés  $K_i$  difféomorphe  $\mathcal{C}^2$  à  $[0, 1]^2$  tels que  $D = \cup K_i$  et  $i \neq j \Rightarrow \mathring{K}_i \cap \mathring{K}_j = \emptyset$ , alors pour tout  $\mathbf{F} : D \rightarrow \mathbb{R}^2$  champ de vecteurs  $\mathcal{C}^1$ ,

$$\int_{\partial D} \mathbf{F} \cdot d\mathbf{s} = \int_D \text{rot } \mathbf{F} d\text{vol}$$

Encore une fois, il ne s'agit pas des conditions optimales de régularité.

Les mêmes méthodes qu'au théorème 9.3.1 (beaucoup de calcul) permettent de montrer que

**Théorème 9.3.4 (Théorème de la divergence, cas d'un morceau)**

Soit  $P \subset \mathbb{R}^3$  un pavé, et soit  $U \supset P$  un ouvert. Soit  $\mathbf{H} : U \rightarrow \mathbb{R}^3$  un difféomorphisme  $\mathcal{C}^2$  de jacobien positif,  $\Phi_{i,\pm} : [0, 1]^2 \rightarrow \mathbf{H}(P_{i,\pm})$  les paramétrages locaux de sorte que  $\mathbf{N}_{\Phi_{i,\pm}}$  est positivement colinéaire à  $\mathbf{N}_{\mathbf{H}|_{P_{i,\pm}}}$ ,  $\mathbf{F} : \text{Im } \mathbf{H} \rightarrow \mathbb{R}^3$  un champ de vecteurs de classe  $\mathcal{C}^1$ . Alors

$$\int_{\text{Im } \mathbf{H}} \text{div } \mathbf{F} d\text{vol} = \sum_{i \in \bar{3}} \left( \int_{\Phi_{i,+}} \mathbf{F} \cdot d\mathbf{N} + \int_{\Phi_{i,-}} \mathbf{F} \cdot d\mathbf{N} \right).$$

Puis, une version pour des domaines un peu plus arbitraire s'obtient par découpage en morceaux.

**Exemple 9.3.5.** Soit  $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  la surface paramétrée définie par  $\Phi(u, v) = \begin{pmatrix} u \\ v \\ u^2 + v^2 \end{pmatrix}$  (c'est le paramétrage "naturel" du graphe de  $F(u, v) = u^2 + v^2$ ). Soit  $\mathbf{G} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  le champ de vecteurs défini par  $\mathbf{G}(x, y, z) = \begin{pmatrix} -y \\ x \\ x^2 + y^2 + z^2 \end{pmatrix}$ . Un calcul simple donne que  $\nabla \wedge \mathbf{G}(x, y, z) = \begin{pmatrix} 2y \\ -2x \\ 2 \end{pmatrix}$ .

Pour connaître le flux de  $\nabla \wedge \mathbf{G}(x, y, z)$  à travers  $\Phi(B_r(\mathbf{0}))$  (l'image par  $\Phi$  de la boule de rayon  $r$  centrée en l'origine), il est possible soit de faire un calcul direct :  $\mathbf{N}_{\Phi}(u, v) = \begin{pmatrix} 1 \\ 0 \\ 2u \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ 2v \end{pmatrix} = \begin{pmatrix} -2u \\ -2v \\ 1 \end{pmatrix}$ , d'où

$$\begin{aligned} \int_{\Phi(B_r(\mathbf{0}))} \text{rot } \mathbf{G} \cdot d\mathbf{N} &= \int_{B_r(\mathbf{0})} \text{rot } \mathbf{G}(\Phi(u, v)) \cdot \mathbf{N}_{\Phi}(u, v) d\text{vol} = \int_{B_r(\mathbf{0})} \begin{pmatrix} 2v \\ -2u \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -2u \\ -2v \\ 1 \end{pmatrix} d\text{vol} \\ &= 2 \int_{B_r(\mathbf{0})} 1 d\text{vol} = 2\pi r^2. \end{aligned}$$

Ou alors, le théorème de Stokes permet de regarder le travail de  $\mathbf{G}$  le long du bord. Ici le bord de  $B_r(\mathbf{0}) \subset \mathbb{R}^2$  est orienté par  $\gamma(t) = \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix}$  où  $t \in [0, 2\pi]$ , ce qui induit l'orientation  $\sigma(t) = \Phi(\gamma(t)) = \begin{pmatrix} r \cos t \\ r \sin t \\ r^2 \end{pmatrix}$  sur le bord de  $\Phi(B_r(\mathbf{0}))$ . D'où

$$\begin{aligned} \int_{\sigma} \mathbf{G} \cdot d\mathbf{s} &= \int_0^{2\pi} \mathbf{G}(\sigma(t)) \cdot \sigma'(t) dt = \int_0^{2\pi} \begin{pmatrix} -r \sin t \\ r \cos t \\ 2r^2 \end{pmatrix} \cdot \begin{pmatrix} -r \sin t \\ r \cos t \\ 0 \end{pmatrix} dt \\ &= r^2 \int_0^{2\pi} dt = 2\pi r^2. \end{aligned}$$

Ce qui donne bien le même résultat. ♣

**Exemple 9.3.6.** divergence ♣

**Remarque 9.3.7.** Ce qui est le plus important à retenir du théorème de Stokes n'est pas tant que c'est une méthode de calcul parfois plus simple d'un flux (ou d'un travail). Mais plutôt que le flux d'un rotationnel ne dépend essentiellement pas de la surface choisie. Plus précisément, toutes les



surfaces (orientables) qui ont le même bord (et qui sont suffisamment régulières) auront le même flux (si elles induisent la même orientation sur le bord). Ceci est particulièrement utile dans la théorie de l'électromagnétisme. ♠

EXERCICE 161: *Un planimètre est un dispositif mécanique qui permet de calculer l'aire d'un domaine du plan tout simplement en parcourant son bord. Le principe est de calculer le travail du champ de vecteurs  $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  défini par  $\mathbf{F}(x, y) = \begin{pmatrix} -y \\ x \end{pmatrix}$ . Soit  $D$  un domaine borné du plan. Poser des conditions sur  $D$ , puis montrer que le travail (du champ de vecteurs  $\mathbf{F}$  le long d'une courbe qui parcourt le bord de  $D$ ) donnera bien l'aire de  $D$ .*

EXERCICE 162: *Soit  $V$  un volume dans  $\mathbb{R}^3$  tel que  $\partial V = S$  est une surface régulière admettant un paramétrage régulier et surjectif hors d'un ensemble négligeable,  $\Phi : D \rightarrow \mathbb{R}^3$ . Montrer que,  $\forall a, b, c \in \mathbb{R}$  tels que  $a + b + c = 1$ ,  $\text{vol}(V) = \int_D G(s, t) \, d\text{vol}(s, t)$  où*

$$G(s, t) = \text{Det} \begin{pmatrix} a\Phi_1(s, t) & b\Phi_2(s, t) & c\Phi_3(s, t) \\ D_1\Phi_1(s, t) & D_1\Phi_2(s, t) & D_1\Phi_3(s, t) \\ D_2\Phi_1(s, t) & D_2\Phi_2(s, t) & D_2\Phi_3(s, t) \end{pmatrix}$$

EXERCICE 163: *Soit  $V$  un volume dans  $\mathbb{R}^3$  tel que  $\partial V = S$  est une nappe régulière admettant un paramétrage régulier et surjectif hors d'un ensemble négligeable,  $\Phi : D \rightarrow \mathbb{R}^3$ . Montrer que,  $\text{vol}(V) = \frac{1}{3} \int_S \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot d\mathbf{N}$ .*

*Étant donné un polyèdre régulier (centré en l'origine), en déduire que, si la distance de l'origine au centre d'une face est 1, le volume est  $\frac{nA}{3}$  où  $A$  est l'aire d'une face et  $n$  le nombre de faces<sup>2</sup>.*

## 9.4 Simple connexité et intégrabilité

Un dernier théorème important montre que la topologie porte un rôle important dans la véracité de "champ à rotationnel nul  $\Rightarrow$  c'est un gradient". Autrement dit, si le domaine de définition des fonctions est fixé, l'image du de l'opérateur gradient diffère du noyau de l'opérateur rotationnel. Cette obstruction topologique est la plus simple incarnation de la cohomologie.

### Théorème 9.4.1

Soit  $U \subset \mathbb{R}^m$  un ouvert simplement connexe. Si  $\mathbf{F} : U \rightarrow \mathbb{R}^m$  est tel que  $\text{rot} \mathbf{F}(\mathbf{x}) = 0$  pour tout  $\mathbf{x} \in U$  alors il existe  $V : U \rightarrow \mathbb{R}$  tel que  $\nabla V = \mathbf{F}$ .

*Parenthèse :* Ici  $m = 2$  ou  $3$ , mais le résultat reste vrai (*mutatis mutandis*) pour des  $m > 3$ .

DÉMONSTRATION: L'idée est de construire  $V$  de la seule manière possible, puis d'utiliser l'annulation du rotationnel et la simple connexité pour montrer que cette définition est conséquente.

2.  $n = 4, 6, 8, 12$  ou  $20$ , pour (respectivement) le tétraèdre (face triangulaires, trois faces se rencontrent en un sommet), le cube (face carré, trois faces se rencontrent en un sommet), l'octaèdre (face triangulaires, quatre faces se rencontrent en un sommet), le dodécaèdre (face pentagonales, trois faces se rencontrent en un sommet) et l'icosaèdre (faces triangulaires, cinq faces se rencontrent en un sommet)

Soit  $\mathbf{x}_0$  fixé. Alors  $V(\mathbf{x}) = \int_{\gamma} \mathbf{F} \cdot d\mathbf{s}$  où  $\gamma : [0, 1] \rightarrow U$  est une courbe  $C^1$  allant de  $\mathbf{x}_0$  à  $\mathbf{x}$ . Pour voir que ceci est bien défini, il faut vérifier que si  $\eta : [0, 1] \rightarrow U$  est une autre telle courbe, alors la valeur de  $V(\mathbf{x})$  obtenue est la même. Soit  $\sigma : [0, 1] \rightarrow U$  définie par

$$\sigma(t) = \begin{cases} \gamma(2t) & \text{si } t \in [0, 1/2]; \\ \eta(2-2t) & \text{si } t \in [1/2, 1]. \end{cases}$$

Alors  $\sigma$  est une courbe fermée. Par simple connexité, elle est réalisée par le bord d'un disque (une surface orientée avec l'orientation induite), ici noté par le paramétrage  $\Phi$ . Ainsi

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{s} - \int_{\eta} \mathbf{F} \cdot d\mathbf{s} = \int_{\sigma} \mathbf{F} \cdot d\mathbf{s} = \int_{\Phi} \text{rot} \mathbf{F} \cdot d\mathbf{N} = 0.$$

Reste à vérifier que  $\nabla V = \mathbf{F}$ . Pour ce faire, il suffit de voir que  $D_i V(\mathbf{x}) = F_i(\mathbf{x})$ . Soit  $\phi : [0, t] \rightarrow U$  la courbe définie par  $\phi(s) = \mathbf{x} + s\mathbf{e}_i$ , alors  $\phi'(s) = \mathbf{e}_i$  et

$$V(\mathbf{x} + t\mathbf{e}_i) - V(\mathbf{x}) = \int_{\phi} \mathbf{F} \cdot d\mathbf{s} = \int_0^t F_i(\mathbf{x} + s\mathbf{e}_i) ds =: f(t).$$

Ainsi  $D_i V(\mathbf{x}) = f'(0) = F_i(\mathbf{x})$ . ■

Par les exercices 23 et 43, l'assertion si dessus est optimale. En effet, le champ de vecteur donné par le gradient de  $\mathbf{T}$  étendu par continuité est un champ de vecteur lisse sur  $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ . Cependant,  $\mathbf{T}$  lui-même ne peut pas être étendu par continuité sur  $\{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ et } y \leq 0\}$ . Il est impossible de trouver une fonction continue  $\mathbf{G}$  qui le ferait pour une raison simple : soit  $\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ , alors le travail le long de  $\gamma$  est toujours positif, et ainsi  $\mathbf{G}(\gamma(2\pi)) \neq \mathbf{G}(\gamma(0))$ . Ceci contredirait  $\gamma(0) = \gamma(2\pi)$ .

## Épilogue, alias Chapitre 10

# Le calcul des variations

Ce dernier chapitre est une forme d'apologie de la dérivée directionnelle (ou de Fréchet). En effet, le chapitre 3 a montré que cette manière de procéder amène souvent à calculer la dérivée d'une fonction qui n'est même pas continue. Cependant, la plupart des problèmes naturels se passe dans des espaces de dimension infinie. En voici quelques uns qui ont été touché lors de ce cours :

- Étant donné deux points, quelle est la courbe qui a la propriété qu'un objet qui glisse sans frottement se rend plus rapidement possible d'un point à l'autre ?
- Quelle est la forme d'un pendule dont la période d'oscillation ne dépend pas de l'amplitude ?
- Étant donné deux demi-droites, comment relier leur extrémités de sorte que la variation de courbure soit minimale ?
- Étant donné un paramétrage  $\Phi : D \rightarrow \mathbb{R}^3$  et deux points dans  $D$ , quelle est la courbe  $\gamma : [a, b] \rightarrow D$  qui relie les deux points et telle que la longueur de  $\Phi \circ \gamma$  est minimale ? (*i.e.* quel est le chemin le plus court sur une surface ?)
- Étant donné une courbe dans  $\mathbb{R}^3$ , quelle est la surface qui possède ce bord et dont l'air est minimale ?

Toutes ces questions sont probablement beaucoup plus naturelles au néophyte que la plupart des choses qui ont été abordées jusqu'ici. Pourtant les méthodes ne s'appliquent pas : l'espace de toutes les déformations possibles d'une courbe ou d'une surface est aussi grand que l'espace de toutes les fonctions (non-dénombrable)<sup>1</sup>. Cependant, une chose reste facile à calculer : la dérivée partielle.

En effet, supposons qu'il nous est donné  $F$  une application (*e.g.* courbe paramétrée), et que la quantité à minimiser est  $\mathcal{E}$ . Cette quantité dépendra typiquement de  $F$  (définie sur un fermé  $K$ ) et de ses dérivées. Dans la plupart des cas il y a des conditions au bord à vérifier :  $F$  sur le bord de  $K$  doit prendre certaines valeurs. Alors pour savoir si  $F$  est un maximum ou minimum de  $\mathcal{E}$ , l'idée naturelle est de regarder  $G \mapsto \mathcal{E}(F + G)$  où  $G$  est une autre application telles que  $G$  est nulle sur le bord de  $K$ . De la sorte,  $F + G$  satisfait les conditions au bord. De plus, avec un peu de chance, la limite

$$\lim_{t \rightarrow 0} \frac{\mathcal{E}(F + tG) - \mathcal{E}(F)}{t}$$

---

1. Bien sûr, avec un peu d'analyse fonctionnelle, il deviendra plus facile de parler de la dérivée de Gâteaux.

a un sens et donnera la dérivée directionnelle de  $F$  dans la direction de  $G$ . En effet, c'est la dérivée de la fonction  $e(t) = \mathcal{E}(F + tG)$ . En particulier, si  $F$  est un minimum local, cette dérivée s'annule pour tout  $G$ . Une équation sur  $F$  est alors obtenue, et elle permet de trouver des candidats pour les minima de  $\mathcal{E}$ . Bien sûr, une analyse plus poussée est requise pour montrer que ces points critiques sont bien des minima.

Deux cas particuliers seront traités ici : celui des géodésiques et celui des surfaces minimales. Ces deux cas ont une particularité, "l'application"  $\mathcal{E}$  a la forme

$$\mathcal{E}(F) = \int_D L(F, F')$$

où  $D$  est un intervalle ou un pavé.

## 10.1 Géodésiques et longueur minimale

Soit fixé  $\Phi : D \rightarrow \mathbb{R}^m$  un paramétrage  $\mathcal{C}^2$  d'une surface. L'objectif est de trouver le chemin le plus court sur la surface entre deux points  $\Phi(\mathbf{p})$  et  $\Phi(\mathbf{q})$ . Autrement dit

$$\mathcal{E}(\gamma) = \int_a^b \|(\Phi \circ \gamma)'(s)\| ds$$

où  $\gamma : [a, b] \rightarrow D$  est une courbe paramétrée  $\mathcal{C}^2$  avec  $\gamma(a) = \mathbf{p}$  et  $\gamma(b) = \mathbf{q}$ . Pour simplifier les calculs, c'est plutôt la fonction

$$\mathcal{E}(\gamma) = \int_a^b \|(\Phi \circ \gamma)'(s)\|^2 ds$$

dont les points critiques seront cherchés. Un argument (qui est ici balayé en fin de section) permet de montrer que les points critiques de ces deux "applications" sont les mêmes (mais la seconde est plus régulière, car sans racine carrée).

L'ensemble des déformations  $\mathcal{C}^1$  de  $\gamma$  est donné par

$$U = \{\sigma : [a, b] \rightarrow \mathbb{R}^2 \text{ de classe } \mathcal{C}^1 \mid \sigma(a) = \sigma(b) = 0\}.$$

En effet, pour  $t$  assez petit l'image de  $\gamma + t\sigma$  sera toujours dans  $D$ . Ainsi, si  $\gamma$  est un point critique : pour tout  $\sigma \in U$ ,

$$e'(t) = \frac{d}{dt} \int_a^b \|(\Phi \circ (\gamma + t\sigma))'(s)\|^2 ds.$$

Pour quand même traiter les choses dans une certaine généralité, il est de bon bon de poser

$$L(\gamma, \gamma') = \|(\Phi \circ \gamma)'(s)\|^2 = \gamma'(s)^T Q_{\Phi, \gamma(s)} \gamma'(s).$$

Puis, en interchangeant l'ordre de la dérivation et de l'intégration (c'est possible car toutes les fonctions et leurs dérivées sont continues et le domaine d'intégration est compact),

$$e'(0) = \int_a^b \frac{d}{dt} L(\gamma + t\sigma, \gamma' + t\sigma') ds = \int_a^b \sum_i \left( \frac{\partial L}{\partial \gamma_i}(\gamma, \gamma') \sigma_i + \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') \sigma'_i \right) ds$$

Ce qui est embêtant à cet étape, c'est que  $\sigma$  et  $\sigma'$  apparaissent alors qu'elles sont reliées de manière peu claire. Une petite astuce-clef est l'intégration par partie. Ici cela s'utilise en regardant

$$\int_a^b \frac{d}{ds} \left( \sigma_i \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') \right) ds$$

qui vaut 0 car  $\sigma(a) = \sigma(b) = 0$ . Mais, en utilisant la règle de Leibniz (*i.e.* de dérivation des produits) :

$$0 = \int_a^b \sigma_i \frac{d}{ds} \left( \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') \right) ds + \int_a^b \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') \sigma'_i ds.$$

Ainsi :

$$e'(0) = \int_a^b \sum_i \sigma_i \left( \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') - \frac{d}{ds} \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') \right)$$

### **Théorème 10.1.1 (Équations d'Euler-Lagrange)**

Soit  $\gamma$  un point critique de  $\mathcal{E}(\gamma) = \int_a^b L(\gamma, \gamma') ds$ , alors  $\forall i$

$$\frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') - \frac{d}{ds} \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') = 0$$

DÉMONSTRATION: Il ne reste plus qu'à montrer que si  $h$  est une fonction telle que  $\int_a^b h(s)f(s)ds = 0$  pour tout  $f$  de classe  $C^2$  sur  $[a, b]$  et nulle au bord, alors  $h \equiv 0$ . Si  $h$  n'est pas identiquement nulle, il existe un point  $x_0 \in [a, b]$  avec  $h(x_0) > 0$  (ou  $< 0$  mais l'argument est identique). Comme  $h$  est continue, il est même vrai que  $h(s) > 0$  sur un petit intervalle contenant  $x_0$ , disons  $[a_0, b_0]$ . Soit  $f(s) = (s - a_0)^3(b_0 - s)^3$  si  $s \in [a_0, b_0]$  et  $f(s) = 0$  sinon (*i.e.*  $s \in [a, b] \setminus [a_0, b_0]$ ). Un petit calcul montre que  $f$  est de classe  $C^2$  sur  $[a, b]$  et positive sur  $[a_0, b_0]$ . Alors  $\int_a^b h(s)f(s)ds > 0$ , une contradiction. ■

En revenant à  $L(\gamma, \gamma') = \gamma'(s)^T Q_{\Phi, \gamma(s)} \gamma'(s)$ , il faut calculer trois dérivées. La première est

$$\begin{aligned} \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') &= \sum_{j,k=1}^2 \gamma'_j \gamma'_k \frac{\partial}{\partial \gamma'_i} D_j \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) \\ &= \sum_{j,k=1}^2 \gamma'_j \gamma'_k \left( D_i D_j \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) + D_j \Phi(\gamma(s)) \cdot D_i D_k \Phi(\gamma(s)) \right) \\ &= 2 \sum_{j,k=1}^2 \gamma'_j \gamma'_k D_i D_j \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) \end{aligned}$$

La seconde est

$$\begin{aligned} \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') &= \sum_{k=1}^2 \gamma'_k D_i \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) + \sum_{j=1}^2 \gamma'_j D_j \Phi(\gamma(s)) \cdot D_i \Phi(\gamma(s)) \\ &= 2 \sum_{k=1}^2 \gamma'_k D_i \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) \end{aligned}$$

Puis,

$$\begin{aligned} \frac{d}{ds} \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') &= 2 \sum_{k=1}^2 \gamma'_k D_i \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) + 2 \sum_{k=1}^2 \sum_{j=1}^2 \gamma'_k \gamma'_j D_j D_i \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) \\ &\quad + 2 \sum_{k=1}^2 \sum_{j=1}^2 \gamma'_k \gamma'_j D_i \Phi(\gamma(s)) \cdot D_j D_k \Phi(\gamma(s)). \end{aligned}$$

L'équation d'Euler-Lagrange est alors :

$$0 = \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') - \frac{d}{ds} \frac{\partial L}{\partial \gamma'_i}(\gamma, \gamma') = 2 \sum_{k=1}^2 \gamma'_k D_i \Phi(\gamma(s)) \cdot D_k \Phi(\gamma(s)) + 2 \sum_{k=1}^2 \sum_{j=1}^2 \gamma'_k \gamma'_j D_i \Phi(\gamma(s)) \cdot D_j D_k \Phi(\gamma(s)).$$

Après ces calculs, la condition (8.5.5) est retrouvée. Autrement dit, les géodésiques sont des points critiques de la longueur. Il est assez facile de se convaincre qu'il ne peut pas y avoir de courbe qui maximise la longueur (il est toujours possible de déformer une courbe pour qu'elle ne soit plus rectifiable, et ainsi que sa longueur soit infinie). De fait, comme déjà remarqué, les géodésiques ne sont pas toujours des minima de la longueur (voir les paragraphes suivant l'exemple 8.5.6). Par contre, une courbe qui minimise la longueur (et qui est au moins  $C^2$ ) est une géodésique.

Un point reste à éclaircir : que  $\sqrt{L}$  et  $L$  donne le même résultat. Un calcul direct n'est pas très convaincant : l'équation de  $\sqrt{L}$  est définitivement plus compliquée. Par contre,  $\int_a^b \sqrt{L}$  est invariant par reparamétrage, ainsi il est possible de supposer que la courbe  $\gamma$  est paramétrée par la longueur d'arc. Ensuite  $\frac{\partial \sqrt{L}}{\partial q} = \frac{1}{2} \sqrt{L} \frac{\partial \sqrt{L}}{\partial q}$ . Mais si  $\gamma$  est paramétrée par la longueur d'arc  $\sqrt{L} = 1$ . Pour finir de se convaincre il faut encore montrer que les solutions de (8.5.5) sont de vitesse constante. Mais ceci a déjà été à lors de la discussion de cette équation. Ce qui montre que les deux problèmes variationnels ont les mêmes points critiques.

Il y a une autre astuce qui marche de manière plus générale pour montrer que les solutions de l'équation d'Euler-Lagrange pour  $L$  auront  $\sqrt{L(\gamma)}$  constant. En effet, comme  $L$  est une fonction homogène de degré  $m = 2$  en  $\gamma'$ ,  $\sum_i \frac{\partial L}{\partial \gamma'_i} \gamma'_i = mL$  (c'est le théorème d'Euler sur les fonctions homogènes ou corollaire 3.4.6). Fort de ceci, si  $\gamma$  satisfait l'équation d'Euler-Lagrange (et si  $L$  ne dépend pas de  $s$  directement, ce qui est aussi le cas) :

$$\begin{aligned} m \frac{d}{ds} L &= \frac{d}{ds} \left( \sum_i \frac{\partial L}{\partial \gamma'_i} \gamma'_i \right) = \sum_i \left( \frac{d}{ds} \frac{\partial L}{\partial \gamma'_i} \right) \gamma'_i + \sum_i \frac{\partial L}{\partial \gamma'_i} \gamma''_i \\ &= \sum_i \frac{\partial L}{\partial \gamma'_i} \gamma'_i + \sum_i \frac{\partial L}{\partial \gamma'_i} \gamma''_i = \frac{d}{ds} L. \end{aligned}$$

Ainsi, si  $m \neq 1$  (pour la longueur  $m = 2$ ), tout point critique aura  $L(\gamma, \gamma')$  constante.

## 10.2 Surfaces minimales et aire minimale

# Bibliographie

Voici quelques livres de la bibliothèque de l'institut de mathématiques qui pourraient servir

Auteur	Titre	Langue	Remarque	Cote
Edwards	Advanced Calculus of several variables	EN	Recouvre bien la matière du cours	G99
do Carmo	Differential Geometry of curves and surfaces	EN	Pour les courbes et les surfaces	M219
Marsden et Tromba	Vector Calculus	EN	Beaucoup de dessins !	B121

La plupart des livres d'introduction se trouvent entre la cote B61 et B123. Le livre de Edwards a aussi l'avantage d'être très bon marché. Certains polycopiés des cours donnés les années précédentes sont disponibles sur Internet.

# Index

- $C^0$ , 37
- $C^1$ , 37
- $C^k$ , 37
- $C^\omega$ , 85
- $C^\infty$ , 37
- $\nabla$ , voir gradient
- $\frac{\partial}{\partial x_i}$ , voir dérivée partielle
- $d\mathbf{F}$ , voir dérivée, 29
- $D_i$ , voir dérivée partielle
- $D_{\mathbf{u}}$ , voir dérivée directionnelle
- $\bar{n}$ , vii
- $\pi_i$ , voir projection canonique
- $P^{en}$ , voir plan tangent à un ensemble de niveau
- $P^{par}$ , voir plan tangent à un paramétrage
- $P^{vit}$ , voir plan tangent des vecteurs vitesse
- $\mathbb{R}^n$ , 1
- $T^{en}$ , voir espace tangent à un ensemble de niveau
- $T^{par}$ , voir espace tangent à un paramétrage
- $T^{vit}$ , voir espace tangent des vecteurs vitesse
- $\mathbf{x}$ , 1
- $\mathbf{0}$ , voir origine
  - $^{-1}$ , voir image réciproque
  - $^{[k]}$ , voir dérivée d'ordre  $k$
- $\circ$ , voir composition
- $\partial$ , voir bord
- $^\circ$ , voir intérieur
- accélération, 76
- adhérence, 15
- admissible
  - ensemble, 143
  - fonction, 139, 143
  - improprement, 152
- aire
  - surface, 171
- angle, 4
- application, 11
  - analytique, 85
  - continue, 19
  - contractante, 101
  - de transition, 160
  - linéaire, 26
- base canonique, 2
- bord, 15
- borné, 16
  - fonction, 138
- boule, 14, 15
- centre de courbure, 78
- champ de vecteurs, 11
  - conservatif, 189
- classe
  - $C^0$ , 37
  - $C^1$ , 37
  - $C^k$ , 37
  - $C^\omega$ , 85
  - $C^\infty$ , 37
- compact, 18
- composition, vii
- connexe par arc, 152
- connexité
  - par arc, 152
- continuité, 19
  - uniforme, 140
- contrainte, 116
- coordonnées, 1
- courbe
  - développée, 78



- équivalente, 56
- fermée, 80
- paramétrée, 11
- rectifiable, 61
- régulière, 58
- simple, 80
- courbure
  - d'une courbe, 77
  - d'une surface, 177
- découpage
  - pavé, 138
- définie
  - négative, 90
  - positive, 90
- degré, 83
  - d'un recouvrement, 134
  - multi-, 83
- dérivabilité, 27
- dérivée, 27
  - composée, 33
  - d'ordre  $k$ , vii
  - de Fréchet, 27
  - de Gâteaux, 54
  - directionnelle, 39
  - partielle, 24
- développement de Taylor, 85
- dimension
  - sous-variété, 113
- discontinuité
  - irréparable, 19
  - réparable, 19
- distance, 14
- division d'intervalle, 59
- élément de longueur, 170
- ensemble
  - admissible, 143
  - affin, 42
  - de niveaux, 12
  - homogène, 46
  - linéaire, 42
  - négligeable, 139
  - symétrique [centralement], 46
- espace
  - tangent
    - au graphe, 48
    - niveau, 43
    - paramétrage, 48
    - vitesse, 46
  - vectorel, 2, 8
- extension
  - par continuité, 20
- extremum, 87
- fermé, 15, 16
- feuille, 137
- flux, 173
- fonction, 11
  - admissible, 139, 143
  - analytique, 85
  - bornée, 138
  - caractéristique, 143
  - continue, 19
  - coordonnée, 11
  - harmonique, 38, 192
  - holomorphe, 193
  - homogène, 35
  - polynomiale, 84
- forme fondamentale
  - première, 170
  - seconde, 177
- forme quadratique, 90
  - du lagrangien, 128
- géodésique, 179
- gradient, 29
- graphe, 12
- harmonique, 38, 192
- hessien[ne], 91
- holomorphe, 193

- image
  - réciproque, vii
- improprement admissible, 152
- Inégalité de Cauchy-Schwarz, 4
- intégrabilité, 138
- intégrable, 138
  - sur un domaine général, 143
- intégrale, 138
  - curviligne, 74
  - de surface, 171
  - impropre, 152
- intérieur, 15
- jacobienne, *voir* matrice de Jacobi
- lagrangien, 118
- largeur
  - d'un découpage, 138
  - d'un rectangle, 137
- limite, 17
- linéarité, 26
- Lipschitz, 61
- longueur
  - courbe paramétrée, 61, 64
  - courbe sur une surface, 170
- matrice, 26
  - de Hesse, 91
  - de Jacobi, 30
- maximum
  - absolu, 21
  - local, 87
- MdMdL, *voir* Multiplicateurs de Lagrange, 119
- mineure, 91
  - principale, 91
- minimum
  - absolu, 21
  - local, 87
- Möbius
  - ruban de, 163
- monotone
  - courbe paramétrée, 63
- mutli-degré, 83
- multiplicateurs de Lagrange, 118
- multiplication scalaire, 1
- négligeable, 139
- non-semi-définie, 90
- norme, 4
- orientabilité, 157
- orientation
  - courbe, 166
  - paramétrage, 166
  - surface, 157
- origine, 2
- orthogonalité, 3, 4
- ouvert, 15, 16
- paramétrage, 12, 56
  - application de transition, 160
  - global, 113
  - local, 111
  - longueur d'arc, 69
  - polaire, 57
- pavé, 137
- plan
  - tangent
    - au graphe, 48
    - niveau, 43
    - paramétrage, 48
    - vitesse, 46
- point
  - d'accumulation, 15
  - de bord, 15
  - critique, 89, 118
  - fixe, 101
  - frontière, 15
  - d'inflexion, 89
  - intérieur, 15
  - isolé, 15
  - limite, 15
  - regulier, 111
  - régulier, 111

- singulier, 50, 89, 110, 119
  - d'une courbe, 59
- polynôme, 84
- potentiel, 189
- première forme fondamentale, 170
- produit
  - extérieur, 5
  - mixte, 6
  - scalaire, 3
  - vectériel, *voir* produit extérieur
- projection canonique, 1
- rayon de courbure
  - d'une courbe, 77
- recouvrement, 134
- rectangle
  - dans  $\mathbb{R}^m$ , 137
- rectifiable, 61
- régulier
  - courbe paramétrée, 58
- reparamétrage, 56
  - surfaces, 156
- ruban de Möbius, 163
- seconde forme fondamentale, 177
- semi-définie
  - négative, 90
  - positive, 90
- somme vectorielle, 1
- sous-variété, 113
  - dimension, 113
- suite, 17
  - convergente, 17
- surface
  - équivalente, 156
  - orientable, 157
  - paramétrée, 11
- symboles de Christoffel, 179
- torsion
  - d'une courbe, 79
- trace
  - d'une courbe paramétrée, 55
- transition, application de, 160
- uniformément
  - continue, 140
- valeur
  - critique, 119
  - régulière, 111
  - singulière, 119
- variété, 113
- vecteur
  - accélération, 76
  - unitaire, 3, 39
  - binormal, 79
  - normal, 77
  - tangent, 76
  - vitesse, 55
- vitesse, 55
- volume
  - d'un ensemble, 143
  - d'un rectangle, 137